DOCUMENT RESUME

ED 110 512 95 TM 004 785

AUTHOR Schrader, W.B.; Hilton, Thomas L.
TITLE Educational Attainment of American High School Seniors in 1960, 1965, and 1972: Feasibility Study. Final Report. Report No. PR-75-13.
INSTITUTION Educational Testing Service, Princeton, N.J.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO ETS-PR-75-13
PUB DATE Apr 75
GRANT NIE-G-74-0050
NOTE 99p.

EDRS PRICE MF-$0.76 HC-$4.43 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; Comparative Analysis; Data Collection; Educational Trends; Equated Scores; *Feasibility Studies; *High School Students; Longitudinal Studies; Mathematics; Occupational Aspiration; Parental Background; Probability; Questionnaires; Reading; Sampling; *Seniors; Student Characteristics; *Surveys; Testing; Test Results
IDENTIFIERS Educational Opportunities Survey; National Longitudinal Study; Project TALENT

ABSTRACT

Project TALENT, the Equality of Educational Opportunity Survey, and the Base-Year Survey of the National Longitudinal Study tested national probability samples of American high school seniors in 1960, 1965, and 1972. This study is concerned with the feasibility of using these data for studying trends in educational attainment. Three separate aspects of the problem were considered. First, a detailed review of sample selection, weighting, and school participation showed that the data would provide a sound basis for comparing performance, and identified certain implications of the existing data for making and interpreting comparisons. Second, comparisons were made of students' high school experiences, their educational and occupational aspirations and plans, and their parents' education and occupation. The pattern of results for the three survey samples, although obscured by differences in questions and by population trends, indicated that comparisons of test performances of the groups would be warranted. Third, a detailed design was prepared for studying the equivalence of selected tests used in the three surveys and for equating test scores on pairs of equivalent tests. Application of these equating results to the test results of the three surveys would provide the desired trend data. (Author)

ED110512

PR-75-13

FINAL REPORT

EDUCATIONAL ATTAINMENT OF AMERICAN HIGH SCHOOL SENIORS IN 1960, 1965, AND 1972: FEASIBILITY STUDY

W. B. Schrader, Principal Investigator

and

Thomas L. Hilton, Co-Investigator

National Institute of Education

Project No. NIE-G-74-0050

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

April 1975
EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY

FINAL REPORT

Educational Attainment of American High School Seniors in 1960, 1965, and 1972: Feasibility Study

National Institute of Education Project No. NIE-G-74-0050

W. B. Schrader, Principal Investigator

and

Thomas L. Hilton, Co-Investigator

Educational Testing Service

Princeton, N. J.

April, 1975

The research reported herein was performed pursuant to a grant with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

ABSTRACT

Project TALENT, the Equality of Educational Opportunity Survey, and the Base-Year Survey of the National Longitudinal Study tested national probability samples of American High School Seniors in 1960, 1965, and 1972. This study is concerned with the feasibility of using these data for studying trends in educational attainment.

Three separate aspects of the problem were considered. First, a detailed review of sample selection, weighting, and school participation showed that the data would provide a sound basis for comparing performance, and identified certain implications of the existing data for making and interpreting comparisons. Second, comparisons were made of students' high school experiences, their educational and occupational aspirations and plans, and their parents' education and occupation. The pattern of results for the three survey samples, although obscured by differences in questions and by population trends, indicated that comparisons of test performances of the groups would be warranted. Third, a detailed design was prepared for studying the equivalence of selected tests used in the three surveys and for equating test scores on pairs of equivalent tests. Application of these equating results to the test results of the three surveys would provide the desired trend data.

Acknowledgments

William H. Angoff, Executive Director, CEEB Programs Division; Albert E. Beaton, Director, Office of Data Analysis Research; John C. Bianchini, Head of Statistical Analysis, Western Office, Thomas F. Donlan, Senior Research Psychologist and Chairman, Measurement Development Research Group; and Frederic M. Lord, Senior Research Psychologist, and Chairman, Psychometric Group were members of an Advisory Committee for this study composed of ETS staff members. Their advice made a substantial contribution to the design of the equating and equivalency study presented in Chapter 3 of this report.

Chapter 1

Sample Design, Weighting, and Data Collection in the 1960, 1965, and 1972 Surveys

Introduction

Project TALENT, The Equality of Educational Opportunity Survey (EEOS), and the Base Year Survey of the National Longitudinal Study (NLS) provide data on the test performance of a national probability sample of American high school seniors. A systematic review of the design and execution of this aspect of the three surveys should be distinctly useful in evaluating the feasibility of using these data for comparing 12th grade students in 1960, 1965, and 1972. This chapter is intended to provide such a review.

Purpose and Scope of the Surveys

Consideration of the purpose and scope of the three surveys is essential to an understanding both of the sampling design and the operational procedures which were used. As summarized in Table 1, all three surveys had very ambitious purposes. It seems fair to say that TALENT emphasized measurement, guidance, and manpower considerations, EEOS gave prominence to minority group questions and to the role of the schools in developing abilities; and NLS was particularly concerned with describing students and schools and with the use of data in long-range planning and research. Despite these differences in emphasis, the measurement of student abilities served as a major focus in each research effort. Table 2 provides a summary of relevant aspects of the scope of each survey. Each of the surveys was a massive enterprise calling for widespread involvement of schools in the collection of large quantities of data about students and schools. Table 2 also provides evidence of significant differences in the allocation of resources for data collection among the three surveys. As will be shown later in this chapter, the smaller number of students in the NLS survey is offset by the precision with which the student sample was defined.

Table 1

Major Purposes

TALENT	EEOS	NLS
To provide relevant information on 7 areas of national concern: 1. Available talent 2. Relationships among aptitudes, interests, and other factors 3. Limiting effects resulting from lack of interest and motivation 4. Factors affecting vocational choice 5. Predictors of creativity and productivity 6. Effectiveness of various types of educational experience 7. Procedures for realizing individual potential	As mandated by Congress, the general purpose was to investigate inequality of educational opportunity in public educational institutions. Specific objectives included: 1. Developing comprehensive statistical information on items considered relevant to school equality, 2. Comparing schools mainly attended by minority students with schools attended mainly by majority students on these indexes, and 3. Studying the relation between characteristics affecting quality of education and student performance on aptitude and achievement tests.	The base-year study of the NLS was envisaged both as providing basic data for a continuing follow-up of students participating in it and as a prototype for surveys of later cohorts of high school seniors. The specific objectives of the base-year survey included the following: 1. Describing high school seniors with respect to a wide range of personal characteristics--attitudes, abilities, interests, plans, knowledge of educational and occupational opportunities, and membership in ethnic and other significant subgroups, 2. Describing the high schools attended by participating students and the characteristics of counselors in those schools, and 3. Developing appropriate data files and access procedures to facilitate the use of the base-year data both for follow-up studies and for other research.

Table 2

General Scope of Survey

TALENT	EEOS	NLS
Main Sample: Five percent of high school seniors (between 400,000 and 500,000 students) Grade Levels: 9-12 Amount of Student Time Needed: Two school days Data Collection Instruments: Completed by student: 23 Aptitude and achievement tests Preference Test Themes Student Activities Inventory Interest Inventory	School Sample: Approximately 700,000 public school students in selected grades, with approximately half of the sample to be white and the other half to be nonwhite. Grade Levels: 1, 3, 6, 9, 12 Amount of Student Time Needed: One school day. Data Collection Instruments: 1. Completed by 12th grade students: 7 aptitude and achievement tests Questionnaire 2. Completed by others: Principal Questionnaire Teacher Questionnaire Superintendent Questionnaire	Desired Sample: 21,600 12th grade students in 1,200 high schools. Grade Level: 12 Amount of Student Time Needed: Approximately three hours. Data Collection Instruments 1. Completed by students: 6 aptitude and achievement tests Student questicnnaire 2. Completed for each student by survey administrator: Student's School Record Information Form 3. Completed by others: School Questionnaire Counselor Questionnaire

Only those phases of each survey which included the national 12th grade sample are considered in this summary.

Characteristics Used in Stratification

A comparison of the basic student and school characteristics used in stratification for selecting the public school sample in each survey yields certain insights on similarity and differences in sampling design. Table 3 shows a summary of the basic stratification characteristics. Although all these surveys used the geographical location of the school in stratification, as would be expected in a national survey, the classifications used were different. (The groupings of states and the District of Columbia used in EEOS and NLS are shown in Table 4.) It is clear that the stratification scheme in TALENT is relatively simple, that EEOS placed heavy stress on the number and percent of nonwhite students in sample selection, and that NLS used a greater variety of characteristics in sample selection. In all three surveys, the selection of a sample of schools was of critical importance in the design for sampling. Table 3 suggests that the data available for sampling are, on the whole, less tightly linked to the significant characteristics of pupils in particular schools than would be optimal for stratification purposes.

Sampling and Weighting in TALENT*

Sampling of Schools

Definition of Sampling Frame. The sample design for the high school sample in Project TALENT called for the selection of approximately five per cent of all senior high schools, and, for schools which did not include a ninth grade, appropriately selected feeder schools. Thus, the desired sampling frame would include all public, private, and parochial schools which included grades 9 through 12. In all, about 26,000 high schools

*This discussion of sampling and weighting is based on a number of Project TALENT publications as follows: Flanagan et al., Designing the Study, 1960; Flanagan et al., Design for a Study of American Youth, 1962; Shaycoft et al., Studies of a Complete Age Group--Age 15, 1963; Flanagan et al., The American High School Student, 1964; and Project TALENT, The Project TALENT Data Bank, 1972. Citations of specific reports will be made only when detailed information is cited from a particular report.

Table 3

Variables and Characteristics Used in Selection of Public School Sample

TALENT	EEOS	NLS
1. Geographical Location: Used 56 strata, as follows: 5 largest cities, District of Columbia, and 50 states (with 5 largest cities removed). 2. Size of senior class 3. Retention ratio (holding-power) of school	1. Geographical Location: Used 7 groups of states and distinguished between counties included in a Standard Metropolitan Statistical Area (SMSA) and counties not included in an SMSA. 2. Percent nonwhite in SMSA or county 3. Estimated number of non-white students in secondary school and in its feeder schools. 4. Percent nonwhite in school	1. Geographical Location: Used 4 groups of states. 2. Size of senior class 3. Percent minority in school or in county in which school is located 4. Proximity to higher education of SMSA or county in which school is located 5. Income class of ZIP Code area or county in which school is located 6. Degree of urbanization of city, SMSA, or county.

Table 4

Comparison of Geographical Regions in EEOS and NLS

EEOS

New England	Mid-Atlantic	Southeast	Southwest	Far West and Rocky Mountain	Great Lakes	Plains
Connecticut	New Jersey	Alabama	Oklahoma	Alaska	Illinois	Iowa
Maine	New York	Arkansas	Texas	California	Indiana	Kansas
Massachusetts	Pennsylvania	Florida	Arizona	Colorado	Michigan	Minnesota
New Hampshire	Delaware	Georgia	New Mexico	Hawaii	Ohio	Missouri
Rhode Island	District of Columbia	Kentucky		Idaho	Wisconsin	Nebraska
Vermont	Maryland	Louisiana		Montana		North Dakota
		Mississippi		Nevada		South Dakota
		North Carolina		Oregon		
		South Carolina		Utah		
		Tennessee		Washington		
		Virginia		Wyoming		
		West Virginia				

NLS

Northeast	South	West	North Central
Connecticut	Delaware	Alabama	Oklahoma	Arizona	Alaska	Illinois	Iowa
Maine	District of Columbia	Arkansas	Texas	New Mexico	California	Indiana	Kansas
Massachusetts	Maryland	Florida			Colorado	Michigan	Minnesota
New Hampshire		Georgia			Hawaii	Ohio	Missouri
Rhode Island		Kentucky			Idaho	Wisconsin	Nebraska
Vermont		Louisiana			Montana		North Dakota
New Jersey		Mississippi			Nevada		South Dakota
New York		North Carolina			Oregon		
Pennsylvania		South Carolina			Utah		
		Tennessee			Washington		
		Virginia			Wyoming		
		West Virginia					

were located. The main list of public schools was obtained from the United States Office of Education, which also provided three supplementary lists and a list of parochial schools. In addition, a list of schools supplied by the Bureau of Internal Revenue included names, but not enrollment data, for some additional schools.

Selection of Stratified Public School Sample. The characteristics used in stratifying the public school sample are shown in Table 5. Only size of senior class was used as the basis for assigning different sampling ratios to different strata. The following sampling ratios were used:

	Sampling Ratio
Public schools with fewer than 25 seniors	1:50
Public schools with 25-99 seniors	1:20
Public schools with 100-399 seniors	1:20
Public schools with 400 or more seniors	1:13

Because it was planned from the outset to test all students in the designated grades, simple random sampling of schools within size strata would have yielded a sample in which each student would have had an equal chance of inclusion. This approach, however, would have selected a large number of small schools and relatively few large schools. Consequently, sampling ratios different from 1 to 20 were used for the largest and smallest schools, and corresponding weights were used in the analysis to compensate for this difference. In selecting schools within the strata thus defined, schools were ordered with respect to the retention ratio, and systematic sampling was used for the final selection.

Exceptions to the basic sampling design were made for the New York and Chicago city school systems. In New York, all junior and senior high schools participated, but only 1 student in 12 was tested. In Chicago, 20 of 38 schools were selected at random, and 1 student in 10 was tested in each selected school. (Nine vocational schools were, however, omitted from the Chicago sample as the result of a misunderstanding.) Sampling within schools rather than sampling schools is advantageous from a theoretical viewpoint, so that the design for these two city systems would yield increased precision.

Table 5

Definition and Sources of Variables and Characteristics Used in Selecting Public School Sample in Project TALENT

Variable	Definition	Source
Geographical Location	5 largest cities District of Columbia 50 states (excluding 5 largest cities)	U. S. Bureau of the Census
Size of senior class	Groupings: (a) Under 25 seniors (b) 25-99 seniors (c) 100-399 seniors (d) 400 or more seniors	U. S. Office of Education
Retention ratio (holding-power) of school	Ratio of number of graduates in 1958-59 school year to number of 10th graders in 1957-58 school year	U. S. Office of Education

Selection of Parochial and Private School Sample. Private and parochial schools were stratified only on the basis of the 56 geographical locations. The sampling ratio was 1:20 for these schools. In all, 125 parochial and 59 private senior high schools were selected for the sample, as compared with 879 public senior high schools.

Sampling of Student Data for Analysis

The enormous number of students (nearly 400,000) included in the TALENT survey made sampling of the data for analysis virtually indispensable. Although a number of different samples were created, it is necessary in this discussion to consider only the sample used for the analysis of the test performance of the 12th grade sample. This sample (designated A-10.0-3 by the authors) included all students having a 6-digit testing number ending in 3 and whose records had been incorporated in the 1963 edition of the master tape file.

The fact that the main analyses for 12th grade students were performed at a time when a substantial number of students were not included in the master tape deserves careful consideration. The practical decision to proceed with the analysis was undoubtedly correct. On the other hand, the effect of the missing data on the results is difficult to assess. After a careful evaluation of this problem, the authors of the 1964 report concluded that "there is no basis for asserting with certainty that norms based on the master tape do not differ significantly from the values that would be obtained if the master tape were complete, but there is considerable reason to believe that any differences that may exist are too slight to have major practical importance." The editing and collating of Project TALENT data continued, fortunately, and there now exists, in the Project TALENT Data Bank, a data source which constitutes a complete record of the tested sample. Comparison of the 1963 master tape sample size with the corresponding sample size in the data bank indicates that about 76% of 12th grade women and 78% of 12th grade men were included in the 1963 master tape. If Project TALENT test data are to be used for year-to-year comparisons with other national probability samples of high school seniors, results based on the complete data bank file can and should be used for such comparisons.

Weighting

Several weighting plans were utilized in different phases of Project TALENT. The weighting which is relevant to this report, however, was designed to reproduce the national population represented by the national high school sample. The weight for each student in a particular school is equal to the reciprocal of the sampling ratio applied in selecting the school divided by the proportion of invited schools in a given stratum which participated in the testing. The first element in the product compensates for the differential selection ratios for different schools. The second element in the product is based on the implicit assumption that the best available estimate of characteristics for students in nonparticipating schools are the characteristics of students enrolled in participating schools in the same stratum. Because a very large proportion of the schools which were invited to participate did so, the effect of this estimation for nonparticipants on the results is small. Because of the variant sampling procedures used in the New York and Chicago city school systems, it was necessary to take account of the sampling of students within schools in arriving at the final weights. The simplicity and clarity of the sampling design and the high degree of participation of the invited schools indicate that the weighting system, although necessary, played a relatively minor role in determing the descriptive statistics on 12th grade students.

Sampling and Weighting in EEOS*

First-Stage Sampling

Definition of Sampling Frame. For the first-stage sampling in EEOS, it was decided to use counties rather than school districts as primary sampling units both because census and other statistics were more readily available for counties and because it was judged that counties had greater internal heterogeneity than school districts. The desirability of internal

*Sampling and weighting for EEOS are discussed in Coleman et al., 1966, 550-554, 558-560, 571-572.

heterogeneity arises because cluster sampling has greater adverse effects on the precision of results for a given sample size when variation between clusters is relatively large and variation within clusters is relatively small. Once the decision to use counties (or the equivalent) was made, the sampling frame for the first-stage sampling was fully defined by current census documents.

Selection of Stratified Sample of SMSAs and Counties. At any given time, every county or county equivalent in the United States can be classified unambiguously as belonging to a Standard Metropolitan Statistical Area (SMSA) or as not belonging to an SMSA. The classification of a county on this basis, however, may change because new SMSAs are created when new combinations of counties meet the standard for inclusion. In the first-stage sampling for EEOS, SMSA was treated as the primary sampling unit for all counties which belonged to an SMSA. At the time the sample was selected, there were 209 SMSAs and 2,674 counties not belonging to an SMSA, as shown in Table 6.

A decisive consideration in the sample selection process was the desire that about half of the pupils in the sample be nonwhite. Because it was also judged necessary that all pupils in the designated grades be included in the sample for any school selected for the sample, it was necessary to stratify the SMSAs and nonmetropolitan counties on the basis of their nonwhite populations and to give greater representation in the first-stage sample to primary sampling units having a large proportion of nonwhites while maintaining broad geographical distribution in both the metropolitan and nonmetropolitan units selected.

The description of the sample selection procedures does not state explicitly all the considerations which determined the sampling ratios for the various strata. The sampling ratios for each stratum, which are shown in the report, make it clear that the estimated size of the nonwhite population in a stratum was a major factor in determining the allocation of the first-stage sample. For purposes of this discussion, the totals for all seven

Table 6

Definition and Sources of Variables and Characteristics Used in Selecting Public School Sample for EEOS

Variable	Definition	Source
Geographical Region	See Table 4	----------
Metropolitan vs. Nonmetropolitan County	Standard Metropolitan Statistical Area (SMSA), as defined by the Bureau of the Census, "consists of a county or group of counties containing at least one city (or twin cities) having a population of 50,000 or more plus adjacent counties which are metropolitan in character and are economically and socially integrated with the central city" (U. S. Bureau of the Census, 1972 c, 135). In some states, various subdivisions equivalent to counties, rather than counties, are the basis for an SMSA. In drawing EEOS sample 209 SMSAs were used, and there were 2,674 counties (or the equivalent) outside of metropolitan areas.	County and City Data Book, 1962, U. S. Bureau of the Census
Estimated percent nonwhite students in SMSA or county	Percent nonwhite was grouped as follows: (a) 70 percent and over (b) 30 to under 70 percent (c) 10 to under 30 percent (d) Under 10 percent. For counties not in an SMSA, the "under 10 percent" category was divided into: (a) Estimated nonwhite enrollment of 100 or more (b) Estimated nonwhite enrollment under 100.	County and City Data Book, 1962. (Other sources are not specified.)

Table 6 (Continued)

Definition and Sources of Variables and Characteristics Used in Selecting Public School Sample for EEOS

Variable	Definition	Source
Estimated number of nonwhite students in secondary schools and in its feeder schools	In order to use enrollment figures for secondary schools having varying grade spans to estimate total grade 1 through 12 enrollment in the school, including feeder schools, grade-span coefficients were calculated using the following method. First, total enrollment in each grade 1 through 12 was computed for each of the seven geographical regions and for Alaska and Hawaii. Then the grade-span coefficient for a given set of grades was computed by dividing total enrollment in grades 1 through 12 by the total enrollment in grades included in the designated grade span. When the enrollment in a particular secondary school is multiplied by the appropriate grade-span coefficient, the result is an estimate of the grade 1 through 12 enrollment for the secondary school and its feeder schools. Using figures for percent nonwhite and enrollment, the number of nonwhite students was calculated.	Statistics of Public Schools, Fall, 1964 (OE-20007-64)

Table 6 (Continued)

Definition and Sources of Variables and Characteristics Used in Selecting Public School Sample for EEOS

Variable	Definition	Source
Percent nonwhite in school	A national inventory of school facilities and personnel conducted by the U. S. Office of Education in the spring of 1962 provided the school list. For each school, total enrollment and grade span was obtained as part of the 1962 survey. This information was listed for each secondary school in 66 SMSAs and 332 counties, and was updated as far as possible using publications of State Departments of Education. Each State Department of Education was asked to update and correct the information for each listed school in its state, and to provide data on the percentage of nonwhite students in each school. The following categories were used for percent nonwhite's within school: (a) 75.1 - 100.0 (b) 50.1 - 75.0 (c) 25.1 - 50.0 (d) 10.1 - 25.0 (e) 0 - 10.0	National Inventory of School Facilities and Personnel (OE-21016)

geographical regions should provide a sufficient basis for describing the first-stage sample selection.

Stratum	Primary Sampling Units: Number in Universe	Primary Sampling Units: Number Selected
SMSAs		
Included in sample with certainty	21	21
Proportion nonwhite:		
30-under 70	13	4
10-under 30	46	17
Under 10	129	28
Nonmetropolitan counties		
Proportion nonwhite		
70 and over	19	10
30-under 70	381	105
10-under 30	368	110
Under 10:		
100 or more nonwhite students	498	55
Less than 100 nonwhite students	1,408	52

The 21 SMSAs which were included in the sample with certainty were those whose estimated total nonwhite enrollment in the designated grades was equal to or greater than the corresponding nonwhite enrollment for the stratum divided by the number of SMSAs to be selected from the stratum.

The actual selection of the sample within each stratum was performed by ordering all elements in the stratum on the basis of estimated number of nonwhite students in the designated grades, and then using systematic sampling with a random start to select the first-stage sample. This is the crucial point which insures that the sample is, in fact, a probability sample. There was one stratum containing 4 SMSAs, none of which was selected for the first-stage sample. This would seem to constitute a minor discrepancy in the sampling procedure.

Second Stage Sampling

The second-stage sampling was based on a substantial amount of information about each secondary school and its feeder schools. These data were obtained mainly from State Departments of Education, as shown in Table 6. As in the first-stage sampling particular emphasis was placed on the number and percent of nonwhite students in determining the number of schools to be selected in a given stratum. Schools in each of the 209 SMSA's and 2,674 nonmetropolitan counties were further stratified according to the percent of nonwhite students, yielding the following categories:

(a) 75.1 - 100.0
(b) 50.1 - 75.0
(c) 25.1 - 50.0
(d) 10.1 - 25.0
(e) 0 - 10.0

The 209 SMSAs included 2,741 high schools, of which 349 were selected for the final sample. The 2,674 nonmetropolitan counties included 1,781 high schools, of which 821 were selected for the final sample. The operation of the second-stage sampling in the nonmetropolitan counties resulted in the inclusion of 467 of the 540 high schools which had 10.0 or more percent nonwhite students, and 354 of the 1,241 high schools which had 10 percent or less nonwhite students.

The geographical distribution of the schools selected for the final sample is as follows:

Geographical Location	Number of Schools
Nonmetropolitan	
Southeast	512
Southwest	125
All other regions	184

Geographical Location	Number of Schools
Metropolitan	
Southeast	89
Southwest	42
New England and Middle Atlantic	97
Great Lakes and Plains	85
Rocky Mountain and Far West	36
Total schools	1,170

Thus, 768 of the 1,170 schools included in the survey samples were from the Southeast or Southwest. For comparison, 310 schools from these two regions were included in the 879 schools selected in the public school sample for Project TALENT.

Sampling of Students for Regression Analysis

Definition of Subgroups. The means and standard deviations of test scores and the proportions to be discussed in Chapter 2 of this report were obtained in connection with the regression analysis for the survey. For this reason, the manner in which the regression analysis samples was determined is of critical importance.

Of the 20 subgroups, 8 included white students, and the remaining four were composed of Mexican American, Puerto Rican, American Indian, and Asian American students. The groupings used for the white and black students were as follows:

Metropolitan

- New England and Middle Atlantic (Northeast)
- Great Lakes and Plains (Midwest)
- Rocky Mountain and Far West (West)
- Southeast
- Southwest

Nonmetropolitan

- Southeast
- Southwest
- All other regions

Thus, the 14 strata based on the distinction between Metropolitan and Nonmetropolitan and 7 geographical regions were reduced to 8 strata for purposes of analysis. The separation of the four minority groups for analysis from the large black and white samples produced a more meaningful differentiation than white-nonwhite. Sample sizes for these groups were presumably not large enough to warrant breakdowns on the basis of geography.

Selection of Samples within Subgroups. The selection of the sample for each of the 20 subgroups was done so that each member of the selected sample would have equal weight within the subgroup. This was accomplished by selecting students in proportion to the weight which would be assigned to his data. Thus, the sample within each subgroup was self-weighting.

Because the weight assigned to a student played an important role in the selection of the subgroup samples, the method for determining these weights needs to be described. The following description is based on a relatively brief description which includes several obvious typographical errors. The weight for each student was defined, as would be expected, by the product of two factors. For purposes of defining the weights, data for schools which belonged to a particular second-stage stratum were merged across SMSAs or counties. Thus, the particular SMSA or county to which a school belonged had no effect on its weight. In addition, the 14 strata based on geographical location were reduced to the 8 used in the formation of the 20-group sample. With these revisions in the stratification structure, every school in the sample belongs to one of 35 first-stage strata (4 for each of the 5 SMSA regions and 5 for each of the 3 nonmetropolitan county regions). It also belongs to one of 175 second-stage strata (5 for each first-stage stratum) depending on its percentage of nonwhite students.

The first factor for any school depends only on the first-stage stratum to which it belongs. Data for this factor were obtained from 1960 census figures on the number of persons of school age in each SMSA or county. The factor equals the number of school age persons in the stratum divided by the number of school age persons in SMSAs or counties selected for the first-stage samples from that stratum. The second factor equals the estimated number of students in grades 1-12 in all schools in the SMSAs or counties selected in the first-stage sample from that stratum divided by

the sum of the estimated number of students in grades 1-12 in all schools which actually participated in the study. The use of participating rather than selected schools in the denominator is intended to adjust for the fact that not all schools chose to participate.

The effect of the simplication of the stratification structure was to reduce greatly the variation of weights across strata. It was possible, with some further grouping of strata, to make the ratio of the largest to the smallest weight in one of the 20 subgroups less than 15 to 1 with only one exception. On the whole, it appears that the merging of counties and SMSAs for the determination of weights was desirable.

Weighting of Subgroups

Because it was decided to include approximately the same number of students (1,000) in each of the 20 self-weighting samples, it was necessary to determine the relative weight for each subgroup for use when data for the various subgroups were pooled. These weights corresponded to "the total size of the population represented by each regional stratum." Unfortunately, the specific procedure by which these weights were obtained is not specified clearly. Variability in the subgroup means is great enough to make any national estimate sensitive to the relative weights given to the various subgroups.

Sampling and Weighting in NLS*

Sampling of Schools for Data Collection

Definition of Sampling Frame. The basic sampling frame for public schools included all schools listed on the United States Office of Education school universe tape which indicated that they had a 12th grade plus those schools which indicated that 11th grade was their highest grade. For 7 states, it was necessary to use listings from the 1969-70 school universe tape for all or part of the listings. Private schools were obtained from a tape supplied by the National Catholic Education Association and from an Office of Education listing of private schools. As with public schools,

---

*Sampling design is discussed in WESTAT, 1972, 1-32, 42-45 and in Hilton et al., 1973, 2-14 to 2-16.

all schools which reported 11th or 12th grade as their highest grade were included. Finally, a listing of area vocational schools was obtained from a current directory of such schools. It was recognized that the sampling frame was incomplete and provision was made in the study design to locate additional schools and incorporate them in the sample (WESTAT, 1972, 4-5).

Formulation of Stratification Plan for Public Schools. As shown in Table 7, an extensive array of stratification variables was utilized in defining the NLS sample.

The first grouping of schools was based on size of 12th grade enrollment, geographical region, and proximity code. Schools in the "less than 300 seniors" group were classified into the four geographical regions and three groups based on proximity codes (shown in Table 7) yielding 12 groups. For the remaining schools, proximity code was divided into only two categories: schools in SMSA's or counties containing the 100 largest central-city school districts vs. all others. For these schools, there were two categories by enrollment size, four categories by geographical region, and two categories by proximity code, yielding 16 further groups. At this point, then, 28 groups of schools had been identified.

For each of these 28 groups of schools, a two-way table was constructed, using percent minority (8 levels) and income class (11 categories).

The 28 tables created in the preceding step provided the basis for dividing all schools into Type A and Type B schools. The classification of schools into those two types was based on a decision to sample schools in communities having family incomes less than $5,000 or having more than 20 percent minority students (Type A Schools) at double the rate assigned to other (Type B) schools. This procedure would yield a final sample of which approximately half the schools would belong to Type A.

The 28 tables included a total of 2,464 (28 x 8 x 11) possible subgroups. These were consolidated into 263 relatively homogeneous major strata. Case was taken that each major stratum contain only Type A or Type B schools. The 263 major strata for public schools

Table 7

Definition and Sources of Variables and Characteristics Used in Selecting Public School Sample for NLS

Variable	Definition	Source
Geographical Region	See Table 4	----------
Size of Senior class in school	Groupings: (a) Less than 300 seniors (b) From 300 to 599 seniors (c) 600 or more seniors Note: for schools listing grade 12 as their highest grade, 11th grade enrollment was used.	U. S. Office of Education school universe tapes.
Percent minority in school or in county in which school is located.	Groupings: (a) Under 5 percent (b) 5-9 percent (c) 10-19 percent (d) 20-39 percent (e) 40-59 percent (f) 60-79 percent (g) 80 percent or over (h) Unknown	Records of the office of Civil Rights (OCR) 1970 census data by counties (used if OCR data not available)
Proximity to higher education of SMSA or county in which school is located	Groupings: (a) SMSAs or counties containing the 100 largest central-city school districts (b) Schools in all other SMSAs or counties which contain a public two-year or four-year college (c) All other schools	U. S. Office of Education higher education universe tape

Table 7 (continued)

Definition and Sources of Variables and Characteristics Used in Selecting Public School Sample for NLS

Variable	Definition	Source
Income class of ZIP code area or county in which school is located	For SMSAs, and counties having 50,000 or larger population, adjusted gross income in 5-digit ZIP code area in which school is located was used. For counties having a 1960 population of less than 50,000, however, ZIP code data were not used. Instead, median county income based on the 1960 census was used, because ZIP code area had been found not to be a satisfactory indicator of income level for schools in such counties. 1960 data were adjusted to make them comparable to the 1966 IRS data. The following 11 categories were used: (a) less than $2,000; (b) $2,000-$2,999; (c) $3,000-$3,999; (d) $4,000-$4,999; (e) $5,000-$5,999; (f) $6,000-$6,999; (g) $7,000-$7,999; (h) $8,000-$8,999; (i) $9,000-$9,999; (j) $10,000-and over; (k) Unknown.	For ZIP code areas, Internal Revenue Service tape based on 1966 individual tax returns. For counties, 1960 census reports
Degree of urbanization of city, SMSA, or county in which school is located	Schools contained in: (a) Ten largest central city school districts (b) Next 90 largest central city school districts (c) Other schools in same SMSA or county as Group (a) (d) Other schools in same SMSA or county as Group (b) (e) All other schools included in SMSAs (f) Nonmetropolitan counties having urban population of 75 percent or more (g) Nonmetropolitan counties having urban population between 50 and 74 percent (h) Nonmetropolitan counties having urban population between 25 and 49 percent (i) Nonmetropolitan counties having urban population less than 25 percent (j) Undetermined	Percent urban for nonmetropolitan counties based on 1970 census data.

were then subdivided into 559 final strata on the basis of the degree of urbanization. Categories for this characteristic are shown in Table 7. Final strata were used both in the selection of the sample and to yield a covariance term for the estimation of sampling error.

Stratification of Parochial and Private Schools. Roman Catholic Schools were stratified on the basis of average parents' income, using the following eight categories:

(1) Less than $7,000
(2) $7,000-$8,999
(3) $9,000-$9,999
(4) $10,000-$11,999
(5) $12,000-$13,999
(6) $14,000-$15,999
(7) $16,000-000
(8) Unknown

These 8 categories and four geographical regions were used in stratifying the Roman Catholic Schools. Only schools having an average parents' income less than $7,000 were classified as Type A schools. The subgroups were combined to yield 22 major strata and 36 final strata for parochial schools.

Private schools were stratified only with respect to geographical region. Four major strata and five final strata were composed of private schools

When all schools are considered, the stratification design yielded 289 major strata and 600 final strata.

Allocation of the Sample to Strata. In determining the number of schools to be selected from each stratum, the initial allocation was based on type of school (A or B) and on size of senior class. For Type A schools and Type B schools separately, allocation was proportional to total senior enrollment, so rounded that an even number of schools

would be assigned to each group. Results were as follows (WESTAT, 1972,21):

	Number of Schools	
Enrollment Size	Type A	Type B
Less than 300	304	286
300-599	188	214
600 or more	108	100
Total	600	600

In the subdivision of major strata into final strata, an effort was to make the size of the group included in each final stratum as uniform as possible, within each of the six groups shown in the preceding table. Once the 600 final strata were created, the initial sample was formed by drawing two schools from each final stratum. For strata composed of schools having senior classes less than 300, selection was proportional to size. For strata composed of larger schools, simple random selection of schools was used. This procedure is equivalent to selecting schools in the larger school sizes in proportion to the average size of the schools in the two size groups rather than in proportion to the actual size of the school. This procedure resulted in an advantageous simplification in the work of sampling with no appreciable loss of precision.

The detailed stratification plan developed for the initial sample facilitated the selection of a sample of replacement schools to be invited to participate if any school initially invited found it impossible to participate. Two replacement schools were selected from each final stratum which included at least four schools. Substantially all final strata met this requirement. To the extent that schools within strata are more homogeneous than schools generally, this procedure reduces the effect on the results of replacing nonparticipating schools.

Sampling of Students within Schools in NLS

From a methodological viewpoint, the provision for sampling of individual students in the NLS study has important advantages over the procedure of testing all students in the designated grade within a school. In particular, it yields greater precision for a given number of students

who are tested or who complete questionnaires. Moreover, it seems likely to require the school to define more sharply which students are in 12th grade, and to permit fuller coverage of the defined group than is true when all students present on the testing day are tested. On the other hand, the effort involved in identifying and assembling a small group of students tends to offset the efficiency gained by testing only a small proportion of students in a particular cluster, and administration of the tests and questionnaires to a fraction of the students may provide less realistic conditions than would be obtained by administration in regular classrooms and homerooms.

In the procedure used by NLS, schools were asked to supply a list of their 12th grade students, excluding adult education students, students who were early graduates no longer in attendance, and any foreign exchange students. If there were more than 18 students in the senior class, the NLS project staff selected a sample of 18 students by simple random sampling, and 5 replacement students were selected at random (provided that enrollment was 23 or greater). Each step involved in preparing the student samples was carefully checked. For practical reasons, in about one-fifth of the schools, sampling was done by telephone rather than by the more formal procedure originally planned. The NLS procedures included a check of the student listing in 59 schools, in connection with a site visit. The error rate in the listing of students was about 1 in 1,000. The experience with the selection of students is discussed at some length in the NLS report (Hilton et al., 1973, 2-16 to 2-22). On the whole, the procedure was judged to be feasible and quite accurate.

Determining Weights for NLS

Because NLS sampled students within schools, the probability that a student will be selected is the product of the probability that a student's school will be selected by the probability that the student will be selected. Account must be taken of the fact that the sampling design permitted replacement of a school in the primary sample by another and of the fact that nonresponse may occur either for a school or for students within the school.

For schools having 18 to 299 students, and for Type A and Type B schools separately, each school was selected with a probability proportional to its estimated twelfth grade enrollment insofar as this could be done in a design which called for extensive stratification. Schools having an estimated 12th grade enrollment less than 18 were treated as if their enrollment were 18. Schools having an enrollment of 300 or greater were stratified into two categories: 300-599 and 600 or greater. Within these two categories, they were treated as equal in size in sampling and in determining the school weights. The decision to treat enrollment sizes less than 18 as 18 permitted the inclusion of very small schools in the sample without giving a very large relative weight to students in any such school which happened to be selected. The decision to assign a uniform size measure within each of two size strata to all schools with senior enrollments of 300 or greater facilitated the operations involved in sample selection and was considered to retain the principal advantages of selecting schools with probability proportional to their size. The probability that a given school will be selected in the primary sample can be obtained by dividing the school's size by the total size of the final stratum to which it belongs, and doubling the result. The doubling is necessary because two schools were selected from each final stratum. Because the appropriate weight for a school is the reciprocal of its probability of selection, the weight for a school in the primary sample equals the size of the final stratum divided by twice the size of the school. These weights were calculated by WESTAT for use in determining the final student weights in the NLS study both for schools in the primary sample and for replacement schools which participated in the study. (WESTAT, 1972, 16-23, 31-32, 42-45).

Because 1,044 rather than 1,200 schools participated in the study, a number of final strata included only one rather than two schools and a few final strata had no participating schools. For the final strata including only one school the WESTAT school weight for that school was doubled. For the few final strata containing no schools, the weight for each final stratum was allocated arbitrarily to other final strata, a process described as "smearing" by the authors (Hilton et al., 1973, Appendix E, E-220 and E-221).

In determining the adjusted student weight which was used in the NLS analysis, a participant was defined as a student for whom one or more of the following types of data were available: a School Record Information Form completed by the coordinator at the school; questionnaire responses, or test scores. Then, for schools which have at least one participant, the adjusted student weight is equal to the adjusted school weight times the number of seniors in the school divided by the number of participants in that school. The use of this procedure adjusts the weights so that the increase in the weight per student compensates for a reduction in the number of students participating. (It may be noted that the average number of participants was approximately 17 per school.) Three additional sets of student weights were calculated for analysis of School Record Information Form, questionnaire data, or test scores separately. (Hilton et al., 1973, Appendix E, E-220 and E-221.)

Participation in the Surveys

The critical importance of securing a high degree of participation both by schools and students was fully recognized in the planning of all three surveys. TALENT invited participation by a mid-November mailing to superintendents for testing in March. TALENT also provided 90 regional coordinators, who were outstanding educators and psychologists in all parts of the United States selected in part on the basis of their ability to work with school people. The role of these coordinators in securing a high degree of participation is acknowledged in the report (Flanagan et al., 1960, III-29). EEOS contacts with the schools began with a letter sent early in June to each chief State School Officer from the U. S. Commissioner of Education, asking that a staff member be appointed to coordinate the program within the state. Each state department of education was asked to supply certain data needed for selecting the sample. As soon as possible after authorization was received, a carefully prepared invitation was sent to the superintendent of each selected school. Some 434 consultants in all parts of the United States were appointed by Educational Testing Service to assist schools in any aspect of the

survey. There was great pressure of time in these activities because the main testing was scheduled for the end of September. (Coleman et al., 1966, 549-550, 554, 556). In NLS, the National Center for Educational Statistics asked the Chief State School Officer in each state and the District of Columbia to appoint a state coordinator, and contacts with the superintendents were made through the state coordinator. When principals were invited to participate, they were informed of the prior approval by the state coordinator and the superintendent. Each principal was asked to appoint a survey administrator to manage the survey activities within the school. An honorarium of $50 was provided for each school or for the school's survey administrator at the discretion of the state department of education and the school administration. A similar sequence of approvals was followed for parochial schools. The telephone was used extensively in securing decisions by schools and in resolving problems arising in participation. As in EEOS, there was great pressure of time. The study began early in February, and school terms ended for nearly all schools by the middle of June (Hilton et al., 1973, 2-6 to 2-12).

Participation in TALENT was quite high, amounting to 92.9 percent of the schools invited to participate. The following table shows the number invited and the percent participating for public, parochial, and private high schools (Flanagan, et al., 1962, 55).

	Number in Sample	Percent of Schools Participating
Public	879	93.5
Parochial	125	91.2
Private	59	86.4
Total	1,063	92.9

Although the percentage of students tested within schools was not determined, there were indications that in many communities, school attendance for the special testing was higher than usual (Flanagan et al., 1962, 167).

The participation rate for schools was noticeably smaller in EEOS than in TALENT. Participation figures showing the percentage of schools which administered both tests and questionnaires to 12th grade students, classified jointly by geographical location and by percent nonwhite, are given in the report (Coleman et al., 1966. 567). The overall participation was 67 percent. By geographical groupings, summary results are as follows:

Geographical Grouping	Number in Sample	Percent of Schools Participating
Nonmetropolitan	512	59
Southeast	125	68
Southwest	184	82
All other regions		
Metropolitan		
Northeast	97	82
Midwest	85	61
Southeast	89	61
Southwest	42	74
West	36	72
Total	1,170	67

When schools were classified by percent nonwhite, results are as follows:

Percent Nonwhite in High School	Number in Sample	Percent of Schools Participating
0.0 - 10.0	516	65
10.1 - 25.0	89	77
25.1 - 50.0	71	73
50.1 - 75.0	25	68
75.1 -100.0	469	65
Total	1,170	67

Although there is some variation in percentage of participation when schools are classified in these two ways, the results suggest that the reasons for participating or not participating are not closely associated with the region in which a school is located or its ethnic composition.

Information on participation in NLS differs from that for the other two surveys in several important ways. First, the study design provided for replacing a school which did not participate with another school from the same final stratum, and for replacing a student who did not participate with another student chosen at random within the same school. Second, because the NLS sample was defined in terms of specific students, percent participation of students was reported. Third, in analyzing nonresponse, it was recognized that nonresponse was different for different components of the survey. For example, nonresponse was noticeably smaller for the School Record Information Form than for the questionnaires and tests. One consequence of these refinements is that it is difficult to formulate the response rates in a form comparable to the other two surveys, although a substantial amount of detailed information is available (Hilton et al., 1973, 4-53 to 4-67, 5-1 to 5-4, Appendix A, A-12 to A-41). The following table provides one basis for considering nonresponse. In this table, a distinction is made between schools in the primary sample and schools in the supplementary sample. The following table, adapted from Table 5-1 of the NLS report, shows the number of students from the primary sample of schools and the total number of students who had data on the tests, on the student questionnaire, and on the School Record Information Form (SRIF) separately, and also shows the number who had data on at least one of the three. The percentages were obtained by dividing these numbers by 21,600, which was the number of students called for by the sampling design.

Kind of Data	Number of Students from Primary Sample	Percent of Desired Number	Number of Students in Total Sample	Percent of Desired Number
Questionnaire	15,563	72	16,409	76
Test	14,962	69	15,625	72
SRIF	16,093	75	17,693	82
Questionnaire, Test, or SRIF	16,126	75	17,726	82

It will be noted that the introduction of replacement schools did not have much effect on the response rate as defined. The table also makes it

clear that, among students in the total sample who had data for at least one of the three kinds, 88 percent had test data and 93 percent had questionnaire data (Hilton et al., 1973, 5-3).

In summary, the participation rate of schools and of students makes it unlikely that the participants in TALENT differed appreciably from the sample as designed. In EEOS and NLS, on the other hand, it cannot safely be assumed that the characteristics of the participants correspond closely to the characteristics of the sample called for by the study design. The extent of nonparticipation in NLS and EEOS does not, of course, imply that the results of these studies are biased. What it does do is to reduce substantially the precision which can be attributed to the results as descriptions of 12th grade students in the United States.

Summary and Conclusions

The three surveys differed markedly in the procedures and characteristics used in defining the samples but all three samples of 12th grade students were well-designed, national probability samples. Moreover, the large number of schools involved in the sample selected for each of the three studies, provide adequately for the statistical inefficiency inherent in the use of the school as a key sampling unit in all three studies. One important difference in the sampling frames, however, must be taken into account in any comparison of test performance of the three groups. EEOS was limited to public school students but TALENT and NLS sampled both public and private school students. Strictly speaking, this would entail developing appropriate descriptive statistics for TALENT and NLS for public school students only.

Weighting presents somewhat more difficult questions. Although the weights used in the three surveys are reasonable, it is conceivable that weights devised specifically for the purpose of estimating test performance in the target population would yield a sufficient gain in precision to justify their use. In particular, the special weights devised for summarizing test data should be used for developing estimates for NLS data. The development of special-purpose weights for use with the EEOS regression subgroups, although not essential, deserves consideration as a way of sharpening the year-to-year comparisons.

Two points arising from program operations have definite implications for year-to-year comparisons. In TALENT, the necessary statistics should be based on the TALENT Data Bank sample rather than the sample analyzed for the 1964 report. Second, allowance needs to be made for growth in ability from early fall, when EEOS tests were administered, to spring, when TALENT and NLS tests were administered. Although the order of magnitude of this growth is probably small (perhaps 5 percent of a standard deviation), it is relevant to the interpretation of year-to-year comparisons.

Finally, the effect of nonparticipation by schools in the EEOS and NLS surveys must be considered. It must be acknowledged that nonparticipation is too great to justify any claim that the year-to-year comparisons would be definitive. On the other hand, it may reasonably be stated that they offer a better basis for comparisons across the years 1960 to 1972 than any alternative data sources.

Chapter 2
Characteristics of High School Seniors in 1960, 1965 and 1972 Samples

Development of the Questionnaires

Student questionnaires constituted an important data source in all three surveys. Both TALENT and EEOS used discrete, multiple choice items only. All twelfth grade students were asked to respond to 374 items and college-bound students were asked to reply to 20 additional items in the TALENT survey. EEOS included a 116-item questionnaire. NLS utilized a more flexible format which permitted grouping of items which were concerned with the same topic and which permitted the use of the same options. In the NLS survey, all 12th grade students were asked to respond to 46 items which called for 198 separate choices and to 9 free-response items. They were also asked to reply to an appropriate subset of the remaining 59 items depending on what they were planning to do during the year following high school graduation.

This chapter is concerned with identifying questions which are sufficiently similar in the three surveys to permit comparison of the students in the three samples with respect to important characteristics. In discussing these topics, attention will be given first, to the student's current activities as a high school student; second, to the student's educational and occupational aspirations and plans; and third, to the important but difficult topic of parents' education and occupation.

Each of the questionnaires was prepared with care. In TALENT, an extensive list of questions was assembled. These questions were screened, rewritten, and checked for clarity in pilot interviews with students. Judgments by distinguished advisory panels and project staff took into account previous research in which the questions had been used (Flanagan et al., 1962, 155). The EEOS report does not describe the specific steps involved in the formulation, writing, review, and revision of the questions used in that survey. Instead, emphasis is placed on the pretesting of the items, in which special effort was made to include both

white and black students and to include lower ability and lower middle class students in the sample. The Grade 12 questionnaire was pretested on 67 students (Coleman et al, 1966, 576-577). The NLS questionnaire was developed initially by the Research Triangle Institute. Both the TALENT and EEOS questionnaires were used, along with other sources, in developing items, and new items were written to cover additional topics. A draft version was administered to 9 high school seniors and revised by USOE before pretesting. A pretest was administered to 727 seniors in 10 states and the responses were evaluated in detail. The version developed by Research Triangle Institute embodied the flexible design and the branching which characterized the final questionnaire (Horvitz et al., 1972, Vol. 1, 23-52, Vol. 2, Appendix H). A revised version of the pretested questionnaire was embodied in the RFP for the base year study. The NLS project staff introduced further revisions designed to reduce ambiguities, to make the language appropriate to high school seniors and inoffensive to minority students, and to embody questions which had been used successfully in other ETS programs. Further revisions were made in response to request from the National Center for Educational Statistics based on requests from the Department of Defense and other user groups within HEW. Finally, the questionnaire was pretested again with high school students (Hilton et al., 1973, 2-33 and 2-34).

Considerations in Interpreting Results

In interpreting TALENT questionnaire results, information provided by the authors (Flanagan et al., 1964, 5-5) on the percentage of students who completed items 1, 100, 150, 200, 250, 300, 350 and 374 is relevant. Data are given for males and females separately and for students of each sex classified into 10 groups on the basis of aptitude level. Students were urged repeatedly in the instructions for the first 6 sections (items 1-374) to "answer each question sincerely and thoughtfully". Dropout was less than 5 percent for all of the 20 groups on items 1, 100, 150, and 200. Even at item 300, only students in the lowest tenth in ability had a dropout rate greater than 10 percent. At item 350, however dropout exceeded 10 percent for all male groups except the top tenth and for the five female groups in the lower half in ability. There was a

noticeable trend for a smaller percentage of students in the lower-scoring groups to respond. On item 374, the highest percent response was 85.3, for females in the second tenth in ability, and the lowest percent response was 47.1, for females in the lowest tenth in ability.

In preparing the summary tables on student characteristics in this report, separate percentages for men and women, reported to one decimal place by TALENT (Flanagan et al., 1964, 5-5 through 5-35), were combined using the weighted N based on all students answering each item. The tabulations were based on all students having complete records in the master tape at the time the analyses were done. Because an appreciable number of students who participated in the survey were not included in the analysis, the results differ to an unknown, but probably small, extent from the results which would have been obtained for all participating students. Although the actual N is not cited specifically for these results, it is clear that over 62,700 students were included in the response tabulations. The percentages were based on weighted N's and nonresponses were allocated on a pro rata basis among the options. For this summary, percent omits have been calculated, using a total weighted N of 1,304,200 as shown in Table 2-1, page 2-6, of the 1964 TALENT report.

The results for EEOS are based on unpublished analyses conducted by Albert E. Beaton as part of the basic analysis for that survey. As described in Chapter 1 of this report, the sample for these analyses included 20 subsamples of approximately 1,000 students each, so selected that each sample would be self weighting. In combining the 20 subsamples to arrive at a national figure, weights based on national population data were employed. It may be useful to compare the proportion of students in various ethnic groups as determined by the EEOS weights and as determined by the weighted student responses in the NLS study. Results are as follows:

	EEOS	NLS
Black	13.9	9.4
Mexican American	2.8	2.7
Puerto Rican	1.0	0.4
Other Latin American	---	0.7

	EEOS	NLS
Oriental	0.9	1.0
American Indian	1.3	1.1
Other	1.1	2.9
Unknown	0.9	---
White	78.1	81.8
	100.0	100.0

Except for the somewhat greater proportion of black students in the EEOS sample, the two samples seem to be in good agreement with respect to ethnic compositions particularly in view of the technical difficulties in obtaining data on this topic.

Directions to students for the EEOS questionnaire differed somewhat from those used in TALENT. They were as follows: "mark the space on the answer sheet corresponding to the answer that is correct for you for each question. Mark only one answer for each question. You may leave out any question you prefer not to answer, but we hope you will answer all of them." (Coleman et al, 1966, p. 644). In the EEOS data discussed in this summary, as in the basic EEOS analysis, "omit" is treated as a separate option. Fewer than 5 percent of students omitted most items under consideration. In a few instances, omits have been prorated among the options to facilitate comparison.

Two respects in which EEOS differed from the other two surveys should also be noted. First, the EEOS sample was limited to public school students. Second, EEOS was administered in the fall and the other two surveys were administered in the spring.

For NLS, an important factor in interpreting the results arises from the fact that in this survey students were permitted or encouraged to regard the completion of any instrument as optional and to feel free to omit any items they might regard as an invasion of privacy. The leaflet inviting students to participate included the statement that "Participation in this study is strictly voluntary". Each section of the questionnaire included the following heading: "Please answer every question unless you are asked to skip to another one. You may

omit any question that you or your parents would consider objectionable". (Hilton et al., 1973, Appendix A.) The emphasis on the student's freedom to omit items was considered essential by the research staff to forestall adverse reactions by students to the survey.

It is difficult to assess the effect of these permissive instructions on the extent to which students omitted items. In the NLS data analysis, provision was made for prorating nonresponses among the options. However, the percentage of students omitting an item is shown in each table. The percentage of "omits", like the percentages for the various options, is based on weighted N's in order to describe the population from which the sample was drawn. The actual number of students who completed questionnaires and who were included in the analysis is 16,409. Because none of the items involved in the branching proved to be useful in comparisons with results of the other two surveys, the sample size for the NLS portion of the tables in this chapter is uniformly 16,409. Results summarized in this report were taken from Appendix B of Hilton et al., 1973.

Student's High School Experiences

Results describing the curriculum or program in which a student is enrolled are shown in Table 8. One striking feature of these results is the similarity in the percentage of students who report being enrolled in an academic or college preparatory program, ranging only from 41.4 in EEOS to 42.9 in NLS. Results for most other high school programs are similar for TALENT and EEOS, although the percentage of students in a commercial or business program is smaller for EEOS (18.6) than for TALENT (22.5). In NLS, the percentage for the general program is 32.9, as compared to 21.8 and 22.3 in the other two programs. Interpretation of this shift is difficult because NLS used a somewhat different format for the question, grouping all programs except general and academic under "vocational or technical". In addition, NLS did not offer a category for other programs than those listed. The NLS analysis plan made it possible to compare the student's self-reported program with his program classification determined by the survey coordinator at the high school from the student's school record. For nearly all students, this classification rather than the student's own report was used in classifying students by program in the main questionnaire analysis. (This analysis

Table 8

High School Curriculum or Program

TALENT		EEOS		NLS	
Which one of the following high school programs or curriculums is most like the one that you are taking? If you have not yet been assigned to a program, which do you expect you will take? (Question 91)		Which one of the following best describes the program or curriculum you are enrolled in? (Question 43)		Which of the following best describes your present high school program? (Question 2)	
%[a]	Response	%	Response	%[a]	Response
21.8	General--a program that does not necessarily prepare you either for college or for work, but in which you take subjects required for graduation and many subjects that you like.	22.3	General	32.9	General
42.8	College Preparatory--a program that gives you the training and credits needed to work toward a regular Bachelor's degree in college.	41.4	College preparatory	42.9	Academic or college preparatory
22.5	Commercial or Business--a program that prepares you to work in an office; for example, as a secretary or bookkeeper.	18.6	Commercial or business		Vocational or Technical:
7.7	Vocational--a program that prepares you to work in a shop or factory, or to enter a trade school, or become an apprentice after high school.	6.6	Vocational	1.6	Agricultural occupations
2.1	Agriculture	1.6	Agriculture	12.0	Business or office occupations
2.9	A program very different from the above.	3.0	Industrial arts	2.5	Distributive education
		4.5	Other	0.9	Health occupations
		2.0	Omit	1.1	Home economics occupations
				6.0	Trade or industrial occupations

[a]Omits: TALENT, 1.3%; NLS, 1.4%.

was based only on students for whom data on sex, program, race, and father's education were available.) The following table was adapted from data in Hilton et al., 1973, Appendix B-1:

	Percent in Each Program	
	Self-report	Coordinator
General	32.1	30.0
Academic	44.1	46.0
Vocational-Technical	23.7	23.9

Of students classified by the coordinator as enrolled in the general program, 60.8 reported enrollment in that program. Of students classified by the coordinator as enrolled in the academic program, 78.8 percent reported enrollment in that program. Of students classified in a vocational-technical program, 67.1 reported enrollment in a vocational-technical program. These results suggest that conceptions of what constitutes a general, academic, or vocational-technical curriculum are not as sharply defined or as uniformly understood as would be desirable from a statistical viewpoint.

Results for student's self-reported grades are shown in Table 9. Here the creativity of the questionnaire authors has gone far to forestall any meaningful comparisons. This particular topic would seem to permit almost limitless variations and improvements. Attempts to resolve the difficulties presented by those variations have not provided a basis for drawing any conclusions from the data. This outcome is particularly regrettable because the data do suggest that fewer EEOS students report A and B grades than is true for the other two surveys. There is no way, however, of judging the effect of asking for an average grade rather than for the student's subjective impression of what his report cards looked like.

Data concerning study hours might be expected to throw some light on similarity of samples, although such comparisons would be confounded with possible trends in student motivation. As it happened, TALENT asked for study hours in school and out-of school per week, EEOS asked for studying outside of school on school days, and NLS asked only about homework. Table 10 shows that TALENT students reported a median of 9.0 hours per week and

Table 9

High School Grades

TALENT

Items 106-113. The following questions ask you to report your grades in courses you have taken in the ninth grade or later. Please consider only semester grades. If you have not taken any courses in the topic, skip the item. In these questions choose the one answer that best describes your grades.

[If your school does not use letter grades, please use the following equivalents:

For a grade of A: Excellent; 90-100
For a grade of B: Good; 80-89
For a grade of C: Average; 70-79
For a grade of D: Fair; 60-69
For a grade below D: Failing; 59 or lower]

My grades in all courses starting with the ninth grade have been: (Question 113)

%[a]	Response
3.8	All A's or equivalent
11.0	Mostly A's or equivalent
30.4	Mostly A's and B's or equivalent
40.7	Mostly B's and C's or equivalent
13.1	Mostly C's and D's or equivalent
1.0	Mostly D's or below or equivalent

EEOS

What is your grade average for all your high school work? (Question 88)

%	Response
7.0	A(either A-, A, or A+)
37.1	B(either B-, B, or B+)
42.7	C(either C-, C, or C+)
3.5	D(either D-, D, or D+)
6.7	Don't know
3.0	Omit

NLS

Which of the following best describes your grades so far in high school? (Question 5)

%[a]	Response
9.2	Mostly A (a numerical average of 90-100)
19.4	About half A and half B (85-89)
20.7	Mostly B (80-84)
28.1	About half B and half C (75-79)
14.6	Mostly C (70-74)
6.8	About half C and half D (65-69)
1.0	Mostly D (60-64)
0.2	Mostly below D (below 60)

[a]Omits: TALENT, 3.8%; NLS, 0.7%.

Table 10

Study Hours

TALENT		EEOS	
One the average, how many hours do you study each week? Include study periods in school as well as studying done at home. (Question 97)		On an average school day, how much time do you spend studying outside of school? (Question 61)	
%[a]	Response	%[a]	Response
2.0	None	8.0	None or almost none
19.9	About 1-4 hours per week	9.3	About 1/2 hour a day
31.4	About 5-9 hours per week	18.4	About 1 hour a day
24.9	About 10-14 hours per week	16.6	About 1 1/2 hours a day
14.3	About 15-19 hours per week	23.7	About 2 hours a day
7.5	About 20 or more hours per week	16.7	About 3 hours a day
		7.3	4 or more hours a day
9.0	Median	1.7	Median

[a]Omits: TALENT, 1.3%; EEOS, 1.7%. Omits for EEOS were prorated.

and EEOS students reported a median of 1.7 hours per school day. Although no conclusions can be drawn about differences in the samples from these data, the results seem to be reasonably consistent for the two groups.

The questions about work for pay were virtually identical in TALENT and EEOS. It is true that TALENT asked about work "during the school year" and EEOS limited its concern to work "during the last year". The NLS questionnaire, which asked about all work whether paid or not, was judged not to be comparable to the other two. Table 11 shows results for TALENT and EEOS. In general, it appears that TALENT students were performing more work for pay during the school year than EEOS students. The median number of hours for the TALENT sample was 2.9; for EEOS, it was 0.7. As it happened, however, the percent working 21 hours or more was 15.7 for EEOS and 13.5 for NLS, reversing the pattern.

Table 12, which is concerned with participation in debating, drama, and music, illustrates the approach of the three surveys to extra-curricular activities. TALENT asked students who participated in an activity to rate their degree of participation on a five-step scale. EEOS used only two degrees of participation (active vs. not very active). NLS asked participants to distinguish between active participation and participation as leaders. The following table combines certain responses to aid in comparing the results:

TALENT	EEOS	NLS
% Responses	% Responses	% Responses
23.1 Extremely active or very active	--- ---	--- ---
--- --- ---	28.7 Active	33.0 Participated actively or leader
14.8 Fairly active	--- ---	--- --- ---
8.3 Not very active or rarely active	9.1 Not very active	--- --- ---
53.7 Not a member	63.6 Not a member	67.1 Have not participated
--- --- ---	1.8 Omit	--- --- ---

Table 11

Work for Pay

TALENT		EEOS	
During the school year, about how many hours a week do you work for pay? Do not include chores done around your own home. (Question 37)		During the last school year about how many hours a week did you work for pay? Do not include chores done around your own home. (Question 89)	
%[a]	Response	%[a]	Response
42.3	None	47.2	None
15.8	About 1-5 hours	14.5	About 1 to 5 hours
11.1	About 6-10 hours	9.6	About 6 to 10 hours
8.5	About 11-15 hours	6.2	About 11 to 15 hours
8.8	About 16-20 hours	6.7	About 16 to 20 hours
13.5	About 21 hours or more	15.7	About 21 hours or more
2.9	Median	0.7	Median

[a]Omits: TALENT, 0.9%; EEOS, 3.3%. Omits for EEOS were prorated.

Table 12

Student's Participation in Debating, Drama, or Music

TALENT		EEOS		NLS	
Organizations Items 1-10. How active have you been in any one or more of the following organizations? Debating, dramatics, or musical clubs or organizations (Question 3)		Did you participate in any debating, dramatics, or musical clubs last year? (Question 72)		Have you participated in any of the following types of activities, either in or out of school this year? (Question 10C) Debating, drama, band, chorus	
%	Response	%	Response	%[a]	Response
10.9	Extremely active	60.3	No	67.1	Have not participated
12.2	Very active	28.7	Yes, I was an active member	26.6	Have participated actively
14.8	Fairly active	5.8	Yes, but I wasn't very active	6.4	Have participated as a leader or officer
4.7	A member, but not very active	3.3	Our school does not have such clubs		
3.6	A member, but rarely active	1.8	Omit		
53.7	Not a member of any of these organizations				

[a]Omits: TALENT, 0.5%; NLS, 2.6%.

If there is a difference in level of participation, the manner in which the questions are formulated seems to have obscured it rather effectively. The trend, if any, appears to be toward reduced participation in the later surveys.

Although a number of questions in each survey dealt with extra-curricular activities, differences in groupings of activities were judged to be too great to warrant detailed comparisons. Questions concerning hobby clubs might be comparable but a detailed comparison of the surveys on this point seemed unlikely to yield useful information.

The surveys paid a good deal of attention to student attitudes. As a result of the different interests and emphasis of the authors, however, virtually no meaningful comparisons could be made between the samples. As shown in Table 13, three statements concerned with values and control of environment were repeated identically in EEOS and NLS. Moreover, the options presented to the students seem reasonably similar, after some grouping.

With respect to the statement: "People who accept their condition of life are happier than those who try to change things," there seems to be distinctly less acceptance of this relatively passive viewpoint in 1972 than in 1965. It seems safe to attribute this difference to a pervasive difference in student's attitudes at the time when the surveys were made as of an idiosyncrasy in the sample for either year.

Two additional statements from the EEOS survey were repeated in the NLS study. Both of these were concerned with control of environment.

With respect to the statement--"Good luck is more important than hard work for success"--The great majority of students (86.5 and 86.3 percent) reject this statement in both samples. However, a somewhat larger percentage (8.8) of NLS than of EEOS (5.8) students express agreement with it. Attitudes toward the other statement on control of environment--"Everytime I try to get ahead, something or somebody stops me"--show a higher percentage of the EEOS students replying "Not sure" (21.3) than of NLS students replying "No opinion" (8.4). Of those who had an opinion, about four fifths of the students disagreed with the statement in both samples.

TABLE 13

Student's Attitudes

Statement	EEOS[a]		NLS[b]	
	%[c]	Response	%[c]	Response
People who accept their condition of life are happier than those who try to change things.	36.6	Agree	31.6	Agree or Strongly Agree
(EEOS: Question 101)	21.9	Not Sure	8.4	No Opinion
(NLS: Question 21G)	41.6	Disagree	60.1	Disagree or Strongly Disagree
	%[d]	Response	%[d]	Response
Good luck is more important than hard work for success.	5.8	Agree	8.8	Agree or Strongly Agree
(EEOS: Question 102)	7.7	Not Sure	5.0	No Opinion
(NLS: Question 21B)	86.5	Disagree	86.3	Disagree or Strongly Disagree
	%[e]	Response	%[e]	Response
Every time I try to get ahead, something or somebody stops me.	16.6	Agree	19.2	Agree or Strongly Agree
(EEOS: Question 103)	21.3	Not Sure	8.4	No Opinion
(NLS: Question 21E)	62.0	Disagree	72.3	Disagree or Strongly Disagree

[a] Omits for EEOS were prorated.

[b] Percentages for the two levels of agreement or disagreement have been combined.

[c] Omits: EEOS, 5.0%; NLS, 1.2%.

[d] Omits: EEOS, 4.9%; NLS, 1.1%.

[e] Omits: EEOS, 5.2%; NLS, 1.4%.

Student's Educational and Occupational Plans

Both EEOS and NLS asked students how much education they "wanted to" (EEOS) or "would like to" (NLS) achieve. As shown in Table 14, the specific options differed in the two surveys. Moreover, the high percentage (31.5) of NLS students who omitted this item raises some question about whether it is reasonable to prorate nonresponses among the options. It seems likely that a higher proportion of students with low aspirations than of students with high aspirations would omit the item. Of the EEOS students, about 60 percent reported that they wanted to achieve at least some college training. In NLS, over 75 percent wanted to attend college. In part, this difference presumably reflects an increasing interest in college attendance between 1965 and 1972. The fact that NLS specifically listed junior college as an option but EEOS did not contribute to the difference, as would the fact that parochial and private school students are included in the NLS but not in the EEOS sample.

Educational expectations (TALENT) and plans (NLS) are compared in Table 15. For this question, as for the preceding one, the percentage of NLS students who omitted the item is relatively high (24.7). Although the percentage of nonresponse is fairly high (7.8) for TALENT also, this high degree of nonresponse is probably attributable in part to the fact that the question occurs late in the questionnaire (item 304). Direct comparison of expectations of some college attendance is prevented by the fact that the TALENT question includes the following option: "I expect to obtain vocational, business school, or junior college training". However, the data do permit obtaining the total percentage of students who expect to attend a four-year college or to go beyond college. It turns out that the percentage is higher for TALENT (52.1) than for NLS (49.7).

Plans for attending college in the year following high school graduation are shown in Table 16. The percentage is: 49.3 for TALENT, 38.3 definitely and 26.9 probably for EEOS, and 49.8 for NLS. The questions on this topic had fairly low nonresponse rates. Although

Table 14

Student's Educational Aspirations

EEOS		NLS	
How far do you want to go in school? (Question 49)		To answer this question, circle one number for the highest level of education you would like to attain. (Question 29A)	
%	Response	%[a]	Response
1.9	I do not want to finish high school	0.7	Less than high school graduation
13.2	I want to finish high school only	5.7	Graduate from high school but not go beyond that
25.8	I want to go to technical, nursing, or business school after high school	19.2	Graduate from high school and then go to a vocational, technical, business, or trade school
11.0	Some college training, but less than 4 years	8.4	Go to a junior college
29.6	I want to graduate from a 4 year college	29.2	Go to a four-year college or university
17.2	I want to do professional or graduate work after I finish college	36.8	Go to a graduate or professional school after college
1.5	Omit		

[a]Omits: 31.5%.

Table 15

Student's Educational Expectations

TALENT		NLS	
What is the greatest amount of education you expect to have during your life? (Question 304)		To answer this question, circle one number for the highest level of education you plan to attain. (Question 29B)	
%[a]	Response	%[a]	Response
1.0	I don't expect to finish high school.	2.2	Less than high school graduation
25.8	I expect to graduate from high school.	17.0	Graduate from high school but not go beyond that
21.0	I expect to obtain vocational, business school, or junior college training.	18.5	Graduate from high school and then go to a vocational, technical, business, or trade school
9.9	I expect to obtain some (less than 4 years) regular college training.	12.6	Go to a junior college
27.1	I expect to graduate from a regular four-year college.	36.9	Go to a four-year college or university
15.1	I expect to study for advanced college degrees.	12.8	Go to a graduate or professional school after college

[a]Omits: TALENT, 7.8%; NLS, 24.7%.

Table 16

College Plans for Year After High School Graduation

TALENT		EEOS		NLS	
When do you plan to start college? (Question 303)		Are you planning to go to college (junior or four year college) next year? (Question 56)		What is the one thing that most likely will take the largest share of your time in the year after you leave high school? (Question 31)	
%[a]	Response	%	Response	%[a]	Response
34.9	I don't plan to go to college.	38.3	Definitely yes	9.1	Taking vocational or technical courses at a trade or business school full-time or part-time
49.3	I plan to start college right after high school.	26.9	Probably yes	10.8	Taking academic courses at a junior or community college full-time or part-time
2.7	I plan to start college after completing military service.	19.6	Probably not	5.4	Taking technical or vocational subjects at a junior or community college full-time or part-time
4.0	I plan to start college after I have worked for a few years.	13.7	Definitely not	33.6	Attending a four-year college or university full-time or part-time
9.1	I may go to college sometime in the future, but my plans are not definite.	1.5	Omit	41.0	(Total for other responses not involving school or college attendance)

[a]Omits: TALENT, 7.4%; NLS, 1.5%.

it is difficult to interpret these results, they are not consistent with an expectation of a rising percentage of students going to college.

Parent's attitudes toward educational goals are described in Table 17. These results show a modest downward trend for the percentage of parents who want only a high school education for children who have reached the senior year of high school, and a modest upward trend in the percentage who want their children to attend professional or graduate school. A college degree or advanced training beyond college was wanted for the students by 50.0 percent of parents in the TALENT sample and by 48.1 percent of fathers and 52.7 percent of mothers in EEOS. At least some college was wanted for the students by 55.0 percent of fathers and 60.7 percent of mothers in the EEOS sample and by 57.7 percent of fathers and 61.1 percent of mothers in the NLS sample. These findings are, if anything, too consistent. Taken at face value, they suggest that parents' attitudes toward having their children obtain a college education showed little change between 1960 and 1972. For the present, no satisfactory hypothesis for accounting for this pattern is apparent.

Results on student's occupational preference for TALENT and NLS are presented in Table 18. Substantial differences in the options offered by the questionnaires complicates the interpretation as does the relatively high nonresponse percentage (22.9) for NLS. The order of responses has been changed to facilitate comparisons. For strict comparability, the fact that TALENT asked students what occupation they would "like to enter" while EEOS asked students what they would "like to do" is troublesome because the term "enter" may suggest an early career stage to some students. In any case, if a relatively broad definition of professional careers is chosen, as was done in grouping career preferences for TALENT in Table 18, 54.4 percent of TALENT students would like to enter a profession as compared to 44.7 percent of NLS students who chose "Professional" to describe their career preference. For technical careers, only 2.0 percent of TALENT students chose the two options presented, while 6.7 percent of NLS students chose the technical category. In TALENT, 12.0 percent chose "Secretary, office clerk, or

Table 17

Parents' Attitudes toward Educational Goals

TALENT		EEOS			NLS		
How much education do your parents or guardians want you to have? (Question 337)		How much education does your father want you to have? (Question 27)			As far as you know, how much schooling do your father and mother (or guardian) want you to get? (Questions 91A and 91B)		
%[a]	Response	Father %	Mother %	Response	Father %[b]	Mother %[b]	Response
1.1	They don't care whether I stay in high school	1.4	1.0	Doesn't care if I finish high school or not	0.2	0.2	Wants me to quit high school without graduating
11.5	High school only	9.9	9.5	Finish high school only	7.0	6.4	Wants me to graduate from high school and stop there
22.2	Vocational school, business school, or junior college	18.3	20.7	Technical, nursing, or business school after high school	20.5	21.6	Wants me to graduate from high school and then go to a vocational, technical, trade, or business school
41.8	A college degree	6.9	8.0	Some college but less than 4 years	9.7	11.0	Wants me to go to a two-year or junior college
8.2	Professional or graduate school	38.4	42.0	Graduate from a 4 year college	36.2	37.5	Wants me to go to a four-year college or university
15.3	I don't know	9.7	10.7	Professional or graduate school	11.8	12.6	Wants me to go to a graduate or professional school after graduating from four-year college or university
		3.8	0.7	Father is not at home	14.5	10.8	I don't know
		9.3	5.8	Don't know			
		2.2	1.6	Omit			

[a]Omits: 15.5%.

[b]Omits: Father, 11.5%; Mother, 10.1%.

Table 18

Student's Occupational Preference

TALENT		NLS	
Which one of the following occupations would you most like to enter? If your choice is not on the list, mark the one that is closest to it. Mark one of these even if you have not definitely made up your mind. (Question 212)		In the column under YOU, circle the one number that goes with the best description of the kind of work you would like to do. The exact job may not be listed but circle the one that comes closest. (Question 25A)	
%[a]	Response[b]	%[a]	Response[b]
3.7	Accountant	44.7	PROFESSIONAL such as accountant, artist, clergyman, dentist, physician, registered nurse, engineer, lawyer, librarian, teacher, writer, scientist, social worker, actor, actress
3.5	Artist or entertainer		
1.3	Biological scientist (biologist, botanist, physiologist, zoologist, etc.)		
1.0	Clergyman (minister, priest, rabbi, etc.)		
0.5	College professor		
1.1	Dentist		
8.9	Engineer (aeronautical, civil, chemical, mechanical, etc.)		
4.9	Elementary school teacher		
1.6	Forester		
5.2	High School teacher		
2.5	Lawyer		
1.1	Mathematician		
5.9	Nurse		
0.9	Pharmacist		
2.2	Physical scientist (chemist, geologist, physicist, astronomer, etc.)		
2.1	Physican		
2.5	Pilot, Airplane		
0.3	Political scientist or economist		
2.2	Social worker		
1.6	Sociologist or psychologist		
1.3	Writer		
0.5	Engineering or scientific aide		
1.5	Medical or dental technician		
2.6	Armed forces officer		
0.7	Enlisted man in the armed forces		
2.9	Businessman	1.8	PROPRIETOR OR OWNER such as owner of a small business, contractor, restaurant owner
0.6	Salesman or saleswoman	3.0	SALES such as salesman, sales clerk, advertising or insurance agent, real estate broker
0.5	Craftsman	7.7	CRAFTSMAN such as baker, automobile mechanic, machinist, painter, plumber, telephone installer, carpenter
3.0	Skilled worker (electrician, machinist, plumber, printer)		
0.7	Structural worker (bricklayer, carpenter, painter, paperhanger, etc.)		
12.0	Secretary, office clerk or typist	2.3	OPERATIVE such as meat cutter; assembler; machine operator; welder; taxicab, bus, or truck driver; gas station attendant
2.7	Barber, beautician	14.7	CLERICAL such as bank teller, bookkeeper, secretary, typist, mail carrier, ticket agent
0.6	Policeman or fireman	4.2	SERVICE such as barber, beautician, practical nurse, private household worker, janitor, waiter
2.0	Farmer	2.2	PROTECTIVE SERVICE such as detective, policeman or guard, sheriff, fireman
5.5	Housewife	1.6	FARMER, FARM MANAGER
9.5	Some other occupation different from any above	2.5	LABORER such as construction worker, car washer, sanitary worker, farm laborer
		3.0	HOMEMAKER OR HOUSEWIFE

[a] Omits: TALENT, 3.5%; NLS, 22.9%.
[b] Order of responses has been changed.

typist", as compared with 14.7 in NLS who chose "CLERICAL such as bank teller, bookkeeper, secretary, typist, mail carrier, ticket agent". On the whole, the results for occupational preference suggest that differences in patterns between the two samples are not large enough to be clearly discernible in the available data.

Similar difficulties of interpretation arise when results presented in Table 19 are considered. In particular, the EEOS question--"When you finish your education, what sort of job do you think you will have?" seems likely to suggest to some students the kind of work they will be doing in their twenties, when they finish school, rather than the work they might expect to be doing in their forties. It should also be noted that "some other occupation" was given by 10.7 percent of TALENT students and "Don't know" was given by 18.0 percent of EEOS students. The fairly substantial difference between TALENT (47.5)and EEOS (37.8) may arise in part from the fact that elementary and secondary school teaching, chosen by 11.8 percent of the TALENT sample, was included in the professional group for TALENT. The EEOS question, however, mentioned college professor but not teacher in its examples of professional occupations.

Data for occupational plans, to the extent that they can be compared for the two samples, seem reasonably similar.

Parent and Family Characteristics

Before discussing the characteristics of the parents of students in the three 12th grade samples, it seems desirable to consider briefly some of the factors at work in determining these data. First of all, it should be recognized that a dramatic increase in retention rate of high schools occurred between 1960 and 1966. The number of high school graduates per 1,000 students entering fifth grade eight years earlier rose from 621 to 732 (U.S. Bureau of the Census, 1973, 131). Preliminary figures for 1972 graduates showed an increment to 750. If there had been no shift in parent's educational and occupational status of parents during the period, this factor would be expected to result in a downward trend, particularly between TALENT and EEOS.

Table 19

Student's Occupational Plans

TALENT		EEOS	
In the following list of occupations, mark the one occupation you expect to make your career after you have completed your education. If your choice is not on the list, mark the one that is closest to it. Mark one of these even if you have not definitely made up your mind. (Question 211)		When you finish your education, what sort of a job do you think you will have? (Question 115)	
%[a]	Response[b]	%[a]	Response[b]
3.8	Accountant	37.8	Professional--such as accountant, artist, clergyman, dentist, doctor, engineer, lawyer, librarian, scientist, college professor, social worker, etc.
2.1	Artist or entertainer		
1.0	Biological scientist (biologist, botanist, physiologist, zoologist, etc.)		
0.9	Clergyman (minister, priest, rabbi, etc.)		
0.4	College professor		
1.0	Dentist		
8.6	Engineer (aeronautical, civil, chemical, mechanical, etc.)		
5.2	Elementary school teacher		
1.1	Forester		
6.6	High school teacher		
1.8	Lawyer		
0.9	Mathematician		
5.3	Nurse		
0.8	Pharmacist		
1.8	Physical scientist (chemist, geologist, physicist, astronomer, etc.)		
1.5	Physician		
1.1	Pilot, Airplane		
0.3	Political scientist or economist		
1.5	Social worker		
1.1	Sociologist or psychologist		
0.7	Writer		
0.5	Engineering or scientific aide		
1.5	Medical or dental technician		
3.8	Businessman	7.8	Technical--such as draftsman, surveyor, medical or dental technician, etc.
0.8	Salesman or saleswoman	2.2	Official--such as manufacturer, officer in a large company, banker, government official or inspector, etc.
0.6	Craftsman	4.5	Manager--such as sales manager, store manager, office manager, factory supervisor, etc. Proprietor or owner--such as owner of a small business, wholesaler, retailer, contractor, restaurant owner, etc.
3.4	Skilled worker (electrician, machinist, plumber, printer)	0.8	Salesman--such as real estate or insurance salesman, factory representative, etc.
0.7	Structural worker (bricklayer, carpenter, painter, paperhanger, etc.)	7.8	Skilled worker or foreman--such as baker, carpenter, electrician, enlisted man in the armed forces, mechanic, plumber, plasterer, tailor, foreman in a factory or mine, etc.
1.2	Enlisted man in the armed forces	18.3	Semiskilled worker--such as factory machine operator, bus or cab driver, meat cutter, etc. Clerical worker--such as bankteller, bookkeeper, sales clerk, office clerk, mail carrier, messenger, etc. Service worker--such as barber, waiter, etc. Protective worker--such as policeman, detective, sheriff, fireman, etc.
16.0	Secretary, office clerk or typist		
2.6	Barber, beautician		
0.8	Policeman or fireman		
2.1	Farmer	1.4	Farm or ranch manager or owner
2.4	Armed forces officer	0.3	Farm worker on one or more than one farm
5.5	Housewife	1.1	Workman or laborer--such as factory or mine worker, fisherman, filling station attendant, longshoreman, etc.
10.7	Some other occupation different from any above	18.0	Don't know

[a] Omits: TALENT, 3.6%; EEOS, 8.5%. Omits for EEOS were prorated.
[b] Order of responses has been changed.

It is clearly important, but very difficult,to estimate the effect of trends in educational level of parents during the period covered. Some data on parents' ages are provided by TALENT (Flanagan et al., 1964, 5-22). The median father was 47 years old and the median mother was 43 years old, according to the students. One fourth of the fathers was 53 or older and one fourth of the mothers was 48 or older. At the other end of the distribution, one fourth of the fathers was 42 or under and one fourth of the mothers was 40 or under. Assuming that the same distribution of ages is roughly appropriate for the other two studies, it is possible to calculate the year in which the middle half (25th percentile to 75th percentile) of parents reached age 18 for the three studies, as follows:

	Father	Mother
TALENT	1925 to 1936	1930 to 1938
EEOS	1931 to 1942	1936 to 1944
NLS	1937 to 1948	1942 to 1950

It is clear that many parents in these groups reached high school graduation age during the depression or during World War II years. On the other hand, it is difficult to evaluate the impact of these events on the educational and occupational careers of the parents.

Results for parents' education in the three samples are shown in Table 20. The following table shows the percentage of fathers and mothers who were reported to have finished college or who had taken graduate or professional training beyond college:

	Father	Mother
TALENT	20.1	20.6
EEOS	11.0	7.0
NLS	19.2	11.5

Table 20

Parents' Education

TALENT			EEOS			NLS		
Mark the one answer indicating the highest level of education your father (mother) reached. Mark the one best answer even if you are not sure. (Questions 218 and 219)			How far in school did your father (mother) go? (Questions 19 and 20)			What was the highest educational level each of the following persons completed? If you are not sure, please give your best guess. (Questions 90A and 90B)		
Father %[a]	Mother %[b]	Response	Father %	Mother %	Response	Father %[a]	Mother %[b]	Response
3.3	1.2	None, or some grade school	6.9	4.0	None, or some grade school	3.0	2.0	Doesn't apply
8.3	5.6	Completed grade school	11.7	9.4	Completed grade school	29.2	25.0	Did not complete high (secondary) school
7.3	4.1	Some high school, but did not graduate	21.8	23.3	Some high school, but did not graduate	30.6	42.9	Finished high school or equivalent
16.2	12.8	Graduated from high school	23.7	34.9	Graduated from high school	1.3	2.0	Adult education program
21.6	22.7	Vocational or business school after high school	4.7	6.9	Technical or business school after high school	5.5	6.2	Business or trade school
22.3	32.4	Some junior or regular college, but did not graduate	8.9	7.5	Some college but less than 4 years	11.3	10.4	Some college
5.5	7.4	Graduated from a regular 4-year college	6.9	5.2	Graduated from a 4 year college	10.4	7.4	Finished college (four years)
6.7	6.1	Master's degree	4.1	1.8	Attended graduate or professional school	2.7	1.8	Attended graduate or professional school (for example, law or medical school), but did not attain a graduate or professional degree
6.4	6.1	Some work toward doctorate or professional degree	9.4	5.6	Don't know	6.1	2.3	Obtained a graduate or professional degree (for example, M.A., Ph.D., or M.D.)
1.5	1.0	Completed doctorate or professional degree	1.9	1.6	Omit			
1.1	0.7	I don't know						

[a]Omits: TALENT, 4.2%; NLS, 11.3%.

[b]Omits: TALENT, 5.3%; NLS, 10.5%.

Results are also available for parents who attended but did not complete college, as follows:

	Father	Mother
TALENT	22.3	32.4
EEOS	8.9	7.5
NLS	11.3	10.4

Finally, results can be summarized for parents who were graduated from high school but who did not attend college:

	Father	Mother
TALENT	37.8	35.5
EEOS	38.4	41.8
NLS	37.4	51.1

The trend for EEOS and NLS seem to be reasonable, assuming a generally increasing level of education in the United States. Parents' education, especially the education of mothers for the TALENT sample seems higher than would be expected, even when the complications in interpreting the results are taken into account.

Data on father's occupation are shown in Table 21. Responses have been rearranged to facilitate comparison. On the whole, these results suggest that the samples are reasonably similar. A large number of relatively minor differences in the questions and analysis can be identified, and exact comparisons cannot be made. However, the following summary of percentages gives a rough idea of the outcomes:

Occupational Group	TALENT	EEOS	NLS
Professional	6.1	7.7	14.4
Technical	2.9	2.3	3.0
Official, Manager, Proprietor, Owner	18.3	18.7	20.8
Sales	4.4	4.3	6.0
Skilled worker or Foreman	20.0	22.0	17.8

Table 21

Father's Occupation

TALENT		EEOS		NLS	
Which one of the following comes closest to describing the work of your father (or the male head of your household)? Mark only one answer. If he works on more than one job, mark the one on which he spends most of his time. If he is now out of work, or if he's retired, mark the one that he did last. (Question 206)		What work does your father do? You probably will not find his exact job listed, but check the one that comes closest. If he is now out of work or if he's retired, mark the one that he usually did. Mark only his main job if he works on more than one. (Question 18)		Under FATHER, circle the one number that best describes the work done by your father (or male guardian). The exact job may not be listed but circle the one that comes closest. If either of your parents is out of work, disabled, retired, or deceased, mark the kind of work that he or she used to do. (Question 25B)	
%[a]	Response[b]	%	Response[b]	%[a]	Response[b]
6.1	Professional--such as actor, accountant, artist, clergyman, dentist, engineer, lawyer, librarian, scientist, etc.	7.7	Professional--such as accountant, artist, clergyman, dentist, doctor, engineer, lawyer, librarian, scientist, college professor, social worker, etc.	14.4	PROFESSIONAL such as accountant, artist, clergyman, dentist, physician, registered nurse, engineer, lawyer, librarian, teacher, writer, scientist, social worker, actor, actress
2.9	Technical--such as draftsman, surveyor, medical or dental technician, etc.	2.3	Technical--such as draftsman, surveyor, medical or dental technician, etc.	3.0	TECHNICAL such as draftsman, medical or dental technician, computer programmer
2.4	Official--such as manufacturer, officer in a large company, banker, government official or inspector, etc.	4.0	Official--such as manufacturer, officer in a large company, banker, government official or inspector, etc.	13.8	MANAGER, ADMINISTRATOR such as sales manager, office manager, school administrator, buyer, restaurant manager, government official
7.3	Manager--such as sales manager, store manager, office manager, business manager, factory supervisor, etc.	14.7	Manager--such as sales manager, store manager, office manager, factory supervisor, etc. Proprietor or owner--such as owner of a small business, wholesaler, retailer, contractor, restaurant owner, etc.	7.0	PROPRIETOR OR OWNER such as owner of a small business, contractor, restaurant owner
8.6	Proprietor or owner--such as owner of a small business, wholesaler, retailer, contractor, restaurant owner, etc.			6.0	SALES such as salesman, sales clerk, advertising or insurance agent, real estate broker
4.4	Salesman--such as real estate or insurance salesman, factory representative, etc.	4.3	Salesman--such as real estate or insurance salesman, factory representative, etc.	17.8	CRAFTSMAN such as baker, automobile mechanic, machinist, painter, plumber, telephone installer, carpenter
20.0	Skilled worker or foreman--such as a baker, carpenter, electrician, enlisted man in the armed forces, mechanic, plumber, plasterer, tailor, foreman in a factory or mine (but not on a farm), etc.	22.0	Skilled worker or foreman--such as baker, carpenter, electrician, enlisted man in the armed forces, mechanic, plumber, plasterer, tailor, foreman in a factory or mine, etc.		

Table 21 (Cont.)

Father's Occupation

TALENT		EEOS		NLS	
%[a]	Response[b]	%	Response[b]	%[a]	Response[b]
7.5	Semi-skilled worker--such as factory machine operator, bus or cab driver, meat cutter, etc.	17.5	Semiskilled worker--such as factory machine operator, bus or cab driver, meat cutter, etc.	11.6	OPERATIVE such as meat cutter; assembler; machine operator; welder; taxicab, bus, or truck driver; gas station attendant
2.8	Clerical worker--such as bank teller, bookkeeper, sales clerk, office clerk, mail carrier, messenger, etc.		Clerical worker--such as bankteller, bookkeeper, sales clerk, office clerk, mail carrier, messenger, etc.	2.8	CLERICAL such as bank teller, bookkeeper, secretary, typist, mail carrier, ticket agent
1.1	Service worker--such as barber, beautician, waiter, etc.		Service worker--such as barber, waiter, etc.	2.2	SERVICE such as barber, beautician, practical nurse, private household worker, janitor, waiter
1.6	Protective worker--such as a policeman, detective, sheriff, fireman		Protective worker--such as policeman, detective, sheriff, fireman, etc.	2.5	PROTECTIVE SERVICE such as detective, policeman or guard, sheriff, fireman
8.8	Farm or ranch owner and/or manager	4.7	Farm or ranch manager or owner	5.1	FARMER, FARM MANAGER
0.6	Farm or ranch foreman	2.0	Farm worker on one or more than one farm	11.0	LABORER such as construction worker, car washer, sanitary worker, farm laborer
1.8	Farm or ranch worker	11.4	Workman or laborer--such as factory or mine worker, fisherman, filling station attendant, longshoreman, etc.	2.7	MILITARY such as career officer, enlisted man or woman in the armed forces
16.5	Workman or laborer--such as factory or mine worker, fisherman, filling station attendant, longshoreman, etc.	5.3	Don't know	0.2	HOMEMAKER OR HOUSEWIFE
0.2	Private household worker--such as a servant, butler, etc.	4.2	Omit		
7.0	I don't know.				

[a]Omits: TALENT, 4.6%; NLS, 21.2%.

[b]Order of responses has been changed.

Occupational Group	TALENT	EEOS	NLS
Semi-skilled worker, Clerical worker, Service worker, Protective worker	13.0	17.5	19.1
Farm or ranch manager, owner, foreman, or worker	11.2	6.7	5.1
Workman or laborer	16.5	11.4	11.0
Others and Don't know	7.2	9.5	2.9
Omit	---	4.2	---

The larger percentage in the professional group and the decrease in farm workers over time is reasonable in the light of population shifts, especially if allowance is made for detailed differences in the wording of the questions in the three surveys. The increase in the percentage in the group of occupations including semi-skilled, clerical, service, and protective workers is probably attributable also a general shift in the occupational structure. The smaller percentage of fathers in the skilled worker or foreman group in NLS may have arisen because foreman was omitted from the description of that group in NLS. The larger percentage in sales in the NLS results is probably the result of the fact that sales clerk was included in "sales" in NLS but was classified as "clerical" in the other two surveys. The percentage for "workman or laborer" in the TALENT sample is higher than would be expected for a group of 1960 high school seniors.

Results for mother's occupation are shown in Table 22. If it can be assumed that a mother who does not have a job outside the home is a housewife, the percentage of housewives is 53.4 for the TALENT sample, 50.1 for the EEOS sample, and 55.5 for the NLS sample. The higher figure for NLS may reflect the fact that in TALENT, a student was told not to check housewife if his or her mother had worker for pay in the last three years, and the figure of 50.1 cited for EEOS does not include any of the 15.5 percent of mothers who had a job outside the home at the

Table 22

Mother's Occupation

TALENT		EEOS		NLS	
Which one of the following comes closest to describing the work of your mother (or the female head of your household)? Mark only one answer. If she does housework in addition to outside work, count only the outside work. If she works on more than one job, mark the most important one. If she usually works, but is now out of work, mark the one that she did last. (Question 208)		Does your mother have a job outside your home? (Question 23)		Under MOTHER, circle the one number that best describes the work done by your mother (or female guardian). The exact job may not be listed but circle the one that comes closest. If either of your parents is out of work, disabled, retired, or deceased, mark the kind of work that he or she used to do. (Question 25C)	
%[a]	Response[b]	%	Response	%[a]	Response[b]
53.4	Housewife only; she has not worked for pay in the last three years.	33.1	Yes, full-time	55.5	HOMEMAKER OR HOUSEWIFE
5.5	Professional--such as actress, accountant, artist, dentist, physician, engineer, lawyer, librarian, scientist, etc.	15.5	Yes, part-time	9.1	PROFESSIONAL such as accountant, artist, clergyman, dentist, physician, registered nurse, engineer, lawyer, librarian, teacher, writer, scientist, social worker, actor, actress
0.4	Technical--such as draftsman, medical or dental technician, etc.	50.1	No	0.7	TECHNICAL such as draftsman, medical or dental technician, computer programmer
1.6	Manager--such as sales manager, store manager, office manager, business manager, factory supervisor, etc.	1.3	Omit	1.7	MANAGER, ADMINISTRATOR such as sales manager, office manager, school administrator, buyer, restaurant manager, government official
0.5	Official--such as manufacturer, officer in a large company, banker, government official or inspector, etc.			1.2	PROPRIETOR OR OWNER such as owner of a small business, contractor, restaurant owner
1.6	Proprietor or owner--such as owner of a small business, wholesaler, retailer, restaurant owner, etc.			3.5	SALES such as salesman, sales clerk, advertising or insurance agent, real estate broker
1.3	Sales--such as real estate, life insurance, etc.			0.7	CRAFTSMAN such as baker, automobile mechanic, machinist, painter, plumber, telephone installer, carpenter
1.6	Skilled worker or forewoman--such as baker, inspector, etc.			3.2	OPERATIVE such as meat cutter; assembler; machine operator; welder; taxicab, bus, or truck driver; gas station attendant
5.9	Semi-skilled worker--such as factory machine operator, cab driver, etc.				

Table 27 (Cont.)

Mother's Occupation

TALENT		EEOS	NLS	
%	Response		%	Response
11.5	Clerical worker--such as bookkeeper, secretary, typist, sales clerk, store clerk, etc.		15.9	CLERICAL such as bank teller, bookkeeper, secretary, typist, mail carrier, ticket agent
4.5	Service worker--such as beautician, waitress, etc.		5.7	SERVICE such as barber, beautician, practical nurse, private household worker, janitor, waiter
0.2	Protective worker--such as policewoman, etc.		0.2	PROTECTIVE SERVICE such as detective, policeman or guard, sheriff, fireman
0.3	Farm or ranch owner and/or manager		1.1	FARMER, FARM MANAGER
0.3	Farm or ranch worker		1.3	LABORER such as construction worker, car washer, sanitary worker, farm laborer
3.5	Worker or laborer--such as charwoman, laundry worker, etc.		0.2	MILITARY such as career officer, enlisted man or woman in the armed forces
3.1	Private household worker--such as housekeeper, maid, laundress, etc.			
4.8	I don't know.			

[a]Omits: TALENT, 3.5%; NLS, 18.0%

[b]Order of responses has been changed.

time of the survey. In NLS, on the other hand, the student could judge whether housewife or job constituted his or her mother's main occupation. Comparisons of TALENT with NLS results for jobs outside the home show an increase in percentage in professional employment from 5.5 to 9.1 and for clerical employment from 11.5 to 15.9. In view of shifts in occupational patterns in the American population, and because of various detailed differences in the questions, the data for mother's occupation do not provide a basis for evaluating the representativeness of the samples in the three surveys.

Table 23 provides data on the number of brothers and sisters who dropped out of high school for TALENT and EEOS samples. NLS considered only the oldest brother or sister, and could not be compared with the others. It will be noted that TALENT was concerned with all brothers and sisters who dropped out while EEOS was concerned only with older brothers and sisters who dropped out. Of course, the response "I have no brothers and sisters" cannot be compared with "Have no older brothers or sisters" nor can the response "none" be compared for the two samples, because younger brothers and sisters may not have advanced far enough in school for dropout to be a serious possibility. However, on the assumption that virtually all brothers and sisters who did drop out would also be older than the student in the survey, it is possible to compare the remaining responses. In general, EEOS has a slightly smaller percentage except for "four or more." Thus, the results for the two samples seem reasonable.

Summary and Conclusions

The most salient finding of this review of questions used in the three surveys is the pervasiveness of change in the way questions on essentially the same topic are formulated. No doubt this outcome is a side effect of highly desirable features of sound questionnaire development activities--pretesting, multiple reviews, committee discussions of questions, sensitivity to current concerns, responsiveness to the purposes of the study, need for comparing results with other surveys, and not least the efforts of the questionnaire authors to be both imaginative and realistic in seeking the perfect formulation of each question. This review suggests the desirability of specific attention to the identification of "marker"

Table 23

High School Drop-Out by Brothers and Sisters

TALENT		EEOS	
How many of your bothers or sisters dropped out of high school without graduating? (Question 201)		How many of your older brothers and sisters left high school before finishing? (Question 12)	
%[a]	Response	%	Response
11.2	I have no brothers or sisters.	37.2	Have no older brothers or sisters
71.6	None	44.5	None
9.8	One	9.2	1
3.6	Two	3.5	2
1.8	Three	1.7	3
2.2	Four or more	0.9	4
		0.5	5
		0.3	6
		0.2	7
		0.6	8 or more
		1.3	Omit

[a]Omits: 1.1%.

items on key topics which would be reproduced verbatim. For these items, the burden of proof would fall on the proponents of improvements. Where possible, supplementary items could be used to obtain needed information not adequately defined in the original formulation. It may be added that the questions discussed in this chapter are those most nearly comparable across surveys. The approaches to certain other topics of interest, including subjects studied in high school, foreign language spoken in the home, and the role of various advisors in students' planning, were judged to be too different to warrant detailed discussion in this report.

With respect to the purpose for which this review was undertaken. the results indicate that the differences between the samples may reasonably be attributed to differences in the formulation of the questions, to trends over time, or both. There seems to be no pattern indicating that any of the samples was perceptibly out of line with the other two. Insofar as the results of this review are concerned, then, comparison of test results for the three student groups would yield valuable trend information. In addition, to the extent that differences in questionnaire responses reflect population trends, the results in this chapter may be helpful in interpreting the test results.

Chapter 3

Design for a Study of the Reading and Mathematical Ability of American High School Seniors in 1960, 1965, and 1972

Introduction

In the spring of 1972, a nationwide probability sample of high school seniors was tested as part of the National Longitudinal Study of the High School Class of 1972. A carefully designed probability sample controlled the selection of these students and the sample actually tested corresponded closely to the sample design. The tests administered included 15-minute tests of reading comprehension and of mathematics. Thus, substantial evidence of the performance level of high school seniors on these educationally significant abilities as of the spring of 1972 has been produced. Tests of reading comprehension and of mathematical ability were also administered to national probability samples of high school seniors as part of the Equality of Educational Opportunity survey in the fall of 1965 and the Project TALENT survey in the spring of 1960. Because different tests were employed in each survey, however, the existing data provide no useful information on possible trends in student performances.

That the abilities measured by the reading comprehension and mathematical ability tests included in these surveys are very important educational outcomes is widely accepted. These abilities also play a prominent role in admissions tests for college and for graduate and professional schools. Thus, they serve both as measures of educational attainment and as predictors of future educational performance. Finally, the development of these abilities is not closely bound to a particular educational program but is the result of a wide variety of experiences, both in school and out of school. It is clear that evidence on how well American high school seniors perform on tests of these abilities over a period of years would be a very useful social indicator.

Extraordinary public attention was recently given to a report that scores on the College Board Scholastic Aptitude Test had declined substantially in the past 10 years. Verbal scores declined from 478 in the 1962-63 testing year to 445 in 1972-73; mathematical scores dropped from 502 to 481 during the same period. Because students who take the Scholastic

Aptitude Test are not representative of all high school seniors, it is not clear whether these shifts arise from a decline in educational performance of high school seniors generally or from a change in the composition of the group which takes the SAT. The difficulty in evaluating these results is a dramatic example of the need for sound data on trends in student ability levels.

Purpose

The general goal of the study is to utilize test data from the three carefully designed and executed large-scale surveys of American high school seniors to identify trends in the reading and mathematical abilities.

The specific tasks which need to be accomplished in the proposed study are:

(1) To collect and analyze data to determine whether the reading comprehension and mathematical ability tests of the National Longitudinal Study (NLS) can properly be equated to corresponding tests in Project TALENT and the Equality of Educational Opportunity Survey (EEOS), and if so, to determine equivalent scores for these tests;

(2) To apply the resulting equations to data from each survey, as appropriate; and

(3) To specify and evaluate methodological considerations which need to be taken into account in interpreting the comparisons.

Choice of Tests

In determining the tests to be included in the equating-equivalence study, the goal is to include those tests which have greatest usefulness as indicators of general educational development and which are likely to prove to be parallel and therefore to justify equating. It is considered desirable to limit the study to the essential test pairs in order to minimize demands on school and student time and to permit a thorough study of the tests included.

Six tests were judged to be particularly suitable for inclusion, as follows:

(1) NLS Reading. This 15-minute test included 20 items based on reading passages. The passages were relatively short (100-200 words)

and emphasized straightforward comprehension. Items considered particularly relevant for minority group students were drawn from the Project Access Reading Tests. Analysis of the test data for a sample of 1,955 students who were slightly more able than the total group tested yielded a reliability coefficient of .797, determined using Kuder-Richardson Formula No. 20. The test was judged to be suitable in difficulty for the group. With respect to speededness, it was found that 82 percent of the students completed the test but that about 4 percent did not complete 75 percent of the items, suggesting a slight degree of speededness. The mean biserial correlation of items with total score was .58. On the whole, the statistical characteristics of the test may be considered satisfactory (Hilton et al., 1973, 2-26 and 2-27, Appendix D, 3, 4, and 6).

Evidence on sex differences and differences between racial groups are available for the 12th grade sample (Hilton et al., 1973, Appendix D, D-578).

For sex, statistics are as follows:

	Mean	Standard Deviation	Percent Based on Weighted N	Actual N
Male	9.84	4.98	49.9	6,766
Female	10.05	4.91	50.1	6,926
Total	9.94	4.94	100.0	13,692

In order to facilitate comparisons with group differences observed on other tests, the observed difference in mean raw scores was divided by the standard deviation for the total group. The result is algebraically equivalent to expressing each mean as a standard score for the total group and subtracting. When this procedure was followed, the advantage of the female students expressed in standard score units was .04.

For race, statistics are as follows:

	Mean	Standard Deviation	Percent Based on Weighted N	Actual N
White	10.41	4.77	90.1	11,816
Black	5.67	4.45	9.9	1,876
Total	9.94	4.94	100.0	13,692

The advantage of white students divided by the standard deviation for the total group is 0.96.

(2) TALENT Reading Comprehension. This 30-minute test included 48 items based on reading passages. Topics included social studies, natural science, and literary materials, including poetry (Flanagan et al., 1962, 105-106). Reliability coefficients for this test were obtained by several methods (Flanagan et al., 1964, 2-14, 2-15, 2-51). Because it was judged that the test was slightly speeded, reliability coefficients based on separately timed halves were preferred, even though they were based on experimental rather than final forms of the test. The observed correlation between halves was adjusted in two ways: First, the Spearman-Brown formula was used to adjust for test length, and, second, the coefficient was adjusted for range of talent using data for a ten percent sample of the high school senior group. The reliability coefficient for boys was .855; for girls, it was .825.

Evidence on sex differences is available for a ten percent sample of 12th grade students, as follows (Flanagan et al., 1974, 3-4, 3-5):

	Mean	Standard Deviation	Percent Based on Weighted N	Actual N
Male	33.023	10.483	47.7	2,946
Female	33.555	9.806	52.3	3,302
Total	33.301	10.138	100.0	6,248

The advantage of female students divided by the standard deviation for the total group is 0.05, which corresponds closely to the value of 0.04 for the NLS test.

(3) EEOS Reading is a 35-minute test which includes 35 items. It was taken from the reading test of the Cooperative Sequential Tests of Educational Progress series. As in the other two tests, items were based on reading passages varied as to topics and skills measured. Although data on the reliability of this test are not available, it is possible to obtain an approximate figure. Using data for 100 11th grade students tested as part of the norming program of the Sequential Tests of Educational Progress, and using the Spearman-Brown formula to adjust for test length, the estimated reliability of the EEOS reading test would be .85. Although no formal analysis of speededness was made, a review of item responses showed little evidence of speededness. (Cooperative Test Division, 1957, 10).

Data on scores for black and white students on the EEOS Reading test are presented in the Supplemental Appendix of the EEOS report (Coleman et al., 1966, 59, 73). The actual sample for each group included 1,000 students from each of 8 strata. Results are as follows:

	Mean	Standard Deviation	Percent Based on Weighted N
Black	58.22	17.12	15.1
White	75.64	16.62	84.9
Total	73.00	17.82	100.0

The advantage of white students divided by the standard deviation for the total group is 0.98, which is slightly higher than the corresponding figure of 0.96 based on the NLS Reading test.

(4) NLS Mathematics. Mathematical ability was measured by a test designed to measure basic competance in mathematics while minimizing computation and excluding algebraic, geometric, or trigonometric skills. The item-type used in this test is called quantitative comparison. The student is given two quantities and asked to decide which of the two is larger, to state that they are equal, or to state that insufficient data are provided to justify a decision (Hilton et al., 1973, 2-27, 2-30).

On the whole, the mathematics test seemed to be appropriate for the group (Hilton et al., 1973, Appendix D, 3, 4, 6). Test reliability estimated by Kuder-Richardson Formula No. 20 was .866. With respect to speededness, 85 percent of the students completed the test, but about 4 percent of the students did not complete 75 percent of the items, suggesting a slight degree of speededness. The difficulty level was well suited to the group.

Data on the comparative performance of male and female and of black and white students are available (Hilton et al., 1973, Appendix D, D-582).

For sex, statistics are as follows:

	Mean	Standard Deviation	Percent Based on Weighted N	Actual N
Male	14.05	7.26	49.9	6,766
Female	12.21	7.09	50.1	6,926
Total	13.13	7.23	100.0	13,692

The advantage of male students divided by the standard deviation for the total group is 0.25.

For race, statistics are as follows:

	Mean	Standard Deviation	Percent Based on Weighted N	Actual N
Black	6.38	6.03	9.9	1,876
White	13.87	6.96	90.1	11,816
Total	13.13	7.23	100.0	13,692

The advantage of white students divided by the standard deviation for the total group is 1.04.

(5) TALENT Mathematics. This test, which was administered with a single 50-minute time limit, included three parts (Flanagan et al., 1964, 122-124). Part I (16 items) is concerned with arithmetic reasoning, Part II (24 items) emphasizes concepts and methods studied in 9th grade algebra, but about 40 percent of the items are devoted to mathematical topics usually taught before 9th grade (e.g., fractions and decimals); and Part III (14 items) samples a variety of topics from courses taught in grades 10 to 12 in college preparatory courses, including plane and solid geometry, more advanced topics in algebra, trigonometry, and analytic geometry and calculus. In addition to scores on each part, subtotals of adjacent parts and total score on all three parts were analyzed in the TALENT research.

Reliability coefficients were determined by Kuder-Richardson Formula No. 21 for both part and total scores and by the formula for the reliability of a sum for the subtotals and total. Results, for 12th grade male and female students separately (using the reliability of a sum results for subtotals and totals), were as follows (Flanagan et al., 1964, 2-14, 2-15).

	Reliability for:	
Part	Male Students	Female Students
I	.766	.729
II	.846	.783
I + II	.890	.853
III	.727	.590
II + III	.891	.831
I + II + III	.915	.874

Data on the comparative performance of male and female students on the various scores are as follows (Flanagan et al., 1964, 3-6 and 3-7):

	Male		Female		Total	
Score	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
I	9.81	3.67	8.77	3.54	9.27	3.64
II	12.38	5.63	10.25	4.85	11.27	5.34
I + II	22.18	8.64	19.03	7.59	20.53	8.26
III	4.41	3.05	3.10	2.31	3.73	2.77
II + III	16.79	8.22	13.35	6.60	14.99	7.61
I + II + III	26.58	11.14	22.10	9.23	24.24	10.43
Percent Based on Weighted N	47.7		52.3		100.0	
Actual N	2,946		3,302		6,248	

The advantage of male students divided by the standard deviation of the total group is as follows:

Score	Ratio
I	0.29
II	0.40
I + II	0.38
III	0.47
II + III	0.45
I + II + III	0.43

Of the various scores, only Part I yields a ratio which is reasonably similar to the value of 0.25 obtained for the NLS Mathematics test.

(6) EEOS Mathematics Achievement, like the Reading Comprehension test used in the survey was taken from the Mathematics test of the Cooperative Sequential Tests of Educational Progress (Coleman et al., 1966, 583). It included 25 items to be answered in 35 minutes. The problems in this test emphasize concepts of measurement and geometry and of function and relation. The form which was used for 12th grade students was designed for use in 10th, 11th, and 12th grades.

Using data for 11th grade students tested as part of the norming program of the Sequential Tests of Educational Progress, and using the

Spearman-Brown formula to adjust for test length, the estimated reliability of the EEOS mathematics test would be .72 (Cooperative Test Division, 1957, 10).

Data on scores for black and white students on the EEOS Mathematics Achievement test are presented in the Supplemental Appendix of the EEOS report (Coleman et al., 1966, 59, 73). The actual sample for each group included 1,000 students from each of 8 strata. Results are as follows:

	Mean	Standard Deviation	Percent Based on Weighted N
Black	37.50	18.40	15.1
White	56.82	17.03	84.9
Total	53.90	18.61	100.0

The advantage of white students divided by the standard deviation for the total group is 1.04, which is the same as the corresponding value of 1.04 obtained for the NLS Mathematics test.

The three surveys included a wide variety of tests in addition to the six proposed for inclusion in this study. In particular, the NLS vocabulary test might be related to the sentence completion and synonyms tests, EEOS and to a vocabulary test included in the TALENT survey. The NLS test is very short (5 minutes) and the TALENT test was not separately timed, however. It appears that the added information provided by these tests would be insufficient to warrant their inclusion in the study.

Evaluation of Equivalence

Although available information about the tests selected for inclusion in this study indicates that they are measuring reasonably similar abilities and have reasonably similar statistical characteristics, a more systematic evaluation of the question of equivalence is an essential part of this proposed study. Angoff (1966) has shown that attempts to equate tests which are not parallel are likely to yield results which may be seriously misleading when used in interpreting scores. The dangers are particularly great when the equating results are used in comparing individual students who took different tests. However, the question of parallelism must be given serious consideration in any application of equated scores.

Two approaches to the evaluation of parallelism are included in this study. The first depends essentially upon the correlation between the two tests and their reliability coefficients. Analysis of the data using this approach will be done before the equating analysis is initiated, because certain possible outcomes may affect the equating procedures in specific instances. The second approach calls for score equating of the two tests based on different student groups. Although the usefulness of this approach depends on the insight with which the student subgroups are chosen, it has the advantage of showing the extent to which the equating lines for different tests differ when different student subgroups (e.g., boys and girls) are used for determining the equating line. Data analysis using this approach will be described as part of the equating design. All basic data needed for the first approach are also required for score equating. Use of the second approach, however, requires the administration of a student questionnaire to provide a basis for differentiating the subgroups.

The initial analysis of equivalence will involve a comparison of the correlation between each pair of tests with reliability coefficients for each test using Kuder-Richardson Formula No. 20. Data on the speededness of each test will be obtained in view of the fact that the Kuder-Richardson formula was designed for unspeeded tests.

It is expected that several of the scores yielded by the TALENT mathematics test will be found not to be sufficiently equivalent to NLS mathematics scores to warrant equating. If particular pairs of tests are found to be measuring the same ability, but differ substantially in reliability, true scores rather than observed scores will be used in the equating.

Equating Design

Four equatings are called for in the proposed study:

(a) NLS Reading (15 minutes) and TALENT Reading Comprehension (30 minutes);

(b) NLS Reading (15 minutes) and EEOS Reading (35 minutes);

(c) NLS Mathematics (15 minutes) and TALENT Mathematics (50 minutes); and

(d) NLS Mathematics (15 minutes) and EEOS Mathematics Achievement.

The plan does not provide for the direct equating of EEOS tests to the corresponding TALENT tests, a decision based on the desire to equate the NLS tests to each earlier test as precisely as possible for a given expenditure.

The following are considered to be important characteristics to be embodied in the equating design:

(1) The equating procedures should permit the line of relation between scores to be curvilinear. (This provision is necessary because test specifications, particularly for the distribution of item difficulties, cannot be assumed to be the same for corresponding tests in different surveys.)

(2) Data collected as part of the equating study should permit an evaluation of the equivalence of each pair of tests to be equated and should also permit the calculation of standard errors of equating at various score levels.

(3) The study should be designed so that the standard error of equating is not likely to exceed four percent of the standard deviation in the vicinity of the mean score.

(4) Although it should not be necessary that the equating sample yield precise 12th grade norms, all reasonable steps should be taken to make the equating sample resemble a norms sample, so that the line of relation will be well defined for the national norms group.

(5) Tests should be administered for equating with the same time limits and the same directions used when they were administered in the national surveys.

(6) The total testing time for any participating student should be modest (less than two hours) and the administrative task to be undertaken by any participating school should also be modest.

It is proposed that the equating procedure be an adaptation of Angoff's Design II (Angoff, 1971, 573-576). In this design, the two tests to be equated are administered so that one random half of the equating sample takes the two tests to be equated in one order and the other random half takes the two tests in the reverse order.

This approach has several significant advantages. First, it can be shown that if scores on the two tests are fairly highly correlated, as is

likely to be true for the tests in this study, a substantial gain in precision of equating arises from the fact that the identical group of students take both tests. Second, this design permits the calculation of the correlation between the tests and of the Kuder-Richardson Formula No. 20 reliability of each test. This information will be useful in evaluating the equivalence of the tests to be equated (Lord, 1950; Angoff, 1971).

The possibility that taking the first test may affect performance on the second test presents some theoretical difficulties. Balancing the order of presentation of the two tests does not completely eliminate this problem, because the effects may not be symmetrical between the two orders. Although this problem is minimal when alternate forms of the same test designed to the same specifications are to be equated, it is a matter of greater concern when, as in the proposed study, the tests to be equated differ in length and in item type. In order to minimize carry-over effects between testings, the study design should provide that the two tests to be equated be administered in different school weeks rather than at a single sitting. This provision is likely to be particularly important for the equating of TALENT Mathematics and NLS Mathematics because the TALENT test includes advanced material which may be frustrating to some of the students. The analysis of the equating data should be performed separately for the two orders of administration and the results should be examined for possible evidence of differential effects of the two orders.

In conducting the testing for the equating study, separate samples of schools should be drawn for each of the four equatings and each order of administration. Each school would be asked to test only one student group. This feature of the study design would make the administrative procedure for each school as simple as possible.

In performing the equating operations, it is proposed that for reading and mathematics tests separately, the TALENT or EEOS test be equated directly to the corresponding NLS test. The procedures to be used for equating are drawn from those used in the Anchor Test Study (Bianchini and Loret, 1972, Vol. 1, 138-140, 144-146). The following steps will be performed separately for reading and mathematics tests:

(1) Equate each pair of tests using the equipercentile method. As in the Anchor Test study, distributions will be determined using a

raw score class interval of 1 and will be smoothed using the Tukey-Cureton method before percentiles are determined. If necessary, further smoothing based on judgment will be done. Linear interpolation will be used in determining selected percentiles, and in determining, for each raw score on the NLS test, the corresponding raw score (to one decimal place) on the test being equated to it.

(2) Using the two equating lines arising from the two orders of administration for a given pair of tests, calculate for each raw score on the NLS test the mean of the two estimated values on the test being equated to it.

A design calling for equating NLS directly to TALENT and to EEOS and for equating TALENT to EEOS through NLS was preferred to a somewhat more complex design in which each test would be equated directly to the other two. This preference is based in part on the conception of NLS as a pivotal test both for comparisons with past surveys and with future surveys.

In order to obtain some information about possible effects of the first testing on the second, the scores on each non-NLS test will be transformed to the NLS scale, using the average line based on both orders of administration. Means and standard deviations for NLS scores and the scores obtained by transforming scores on other tests to the NLS raw-score scale will then be compared. Although this procedure will not provide a rigorous evaluation of the effects of the first testing on the second, the results should throw some light on the order of magnitude of the effects.

In addition to the equating based on all students, separate equatings will be performed on designated subgroups of students. If equating results differ substantially from one subgroup to another, after taking account of sampling error, the tests cannot be regarded as parallel.

The particular comparisons to be made will call for dichotomization of the total group on the following five variables:

(1) Sex
(2) Parents' Education
(3) Type of Program in which enrolled (Academic vs. all others)
(4) Self-reported High School Grades
(5) Amount of mathematics study in grades 9 through 12 (Mathematics equating only)

In the three variables calling for a choice of cutting point for dividing the total group into the two subgroups, the cutting point will be chosen so that the two subgroups are as nearly equal in size as the data permit.

Data for forming the designated groups will be obtained from a brief (five-item) questionnaire to be administered along with the NLS test in each pair. Except for the question on mathematics preparation, the appropriate questions from the NLS questionnaire can be used.

Evaluation of Equating Error

Because information on the sampling error of the equating will be of critical importance in applications of the equating results, it is essential to have detailed information on this point. For this purpose, it is proposed to use the method of balanced repeated replication described by McCarthy (1969), and by Kish and Frankel (1970, 1974). Essentially, this method depends on defining a limited number of principal strata and making two independent selections from each stratum. It is then possible to create pairs of non-overlapping half-samples by assigning one of the two selections from each stratum to one half-sample and assigning the other selection from the stratum to the other half-sample. For greatest efficiency in using the available data, a systematic balanced design for the definition of half-samples is desirable. The necessary design principles are provided by Plackett and Burman (1943-46).

An important feature of the balanced half-sample design is that the number of balanced half-samples must be a multiple of four. This restriction does not apply, however, to the number of principal strata. If the number of principal strata is not a multiple of four, then the square matrix for the next larger multiple of four is used. In the design matrix, rows correspond to half-samples and columns to strata. If the number of strata is not a multiple of four, only as many columns of the matrix are used in designing the analysis as there are strata. Thus, if there were nine principal strata, 9 columns and 12 rows of a 12 x 12 design matrix would be used. (McCarthy, 1969).

Although the balanced half-sample design should provide useful information on the extent of equating error, McCarthy (1969, 245) points out that: "the exact characteristics of estimates of variance are for the most part, unknown." Kish and Frankel (1974, 19), on the basis of extensive empirical studies, conclude

that balanced repeated replication designs provide the best available method for estimating sampling error for complex statistics.

The use of this method for a particular equating in the proposed study will require the following steps:

(a) In each of the 11 principal strata designate one-half of the schools as "+" and the other half as "-".

(b) Create 12 half-samples by selecting either the "+" or the "-" schools from each of the 11 principal strata. For example, the first half-sample would be defined as follows: Stratum 1, +; Stratum 2, -; Stratum 3, +; Stratum 4, -; Stratum 5, -; Stratum 6, -; Stratum 7, +; Stratum 8, +; Stratum 9, +; Stratum 10, -; Stratum 11, +. The remaining half-samples would be formed according to the same principles, using the design for 12 half-samples given by Plackett and Burman (1943-46). Corresponding half-samples from each testing order would be formed.

(c) Perform the equating of the two tests using data from each of the 12 half-samples. There would be, accordingly, 12 values of the corresponding score on the EEOS or TALENT for each raw score on the NLS test. In addition, perform the equating based on the entire sample.

(d) The square root of the average of the squared deviations of the 12 half-sample values for a given raw score on NLS from the corresponding score based on the total group may be considered the standard error of equating at that point on the NLS scale.

This procedure for determining the standard error of equating would be used for each of the four equatings included in the study. It may also be used to evaluate the possible role of sampling fluctuations in producing differences in equating lines for different subgroups if a substantial difference in equating lines occurs.

Test Administration

The testing needed for this study requires that eight sequences of tests be administered, as follows

First Test	Time Limit	Scoring Method	Second Test	Time Limit	Scoring Method
TALENT Reading	30'	Rights	NLS Reading	15'	Formula
NLS Reading	15'	Formula	TALENT Reading	30'	Rights
EEOS Reading	35'	Rights	NLS Reading	15'	Formula
NLS Reading	15'	Formula	EEOS Reading	35'	Rights
TALENT Mathematics	50'	Rights	NLS Mathematics	15'	Formula
NLS Mathematics	15'	Formula	TALENT Mathematics	50'	Rights
EEOS Mathematics	35'	Rights	NLS Mathematics	15'	Formula
NLS Mathematics	15'	Formula	EEOS Mathematics	35'	Rights

The foregoing summary makes it clear that in every instance, one number of the pair would require students to respond in terms of "rights-scoring" directions and the other would require students to respond in terms of "formula-scoring" directions. This fact should not present a problem if tests are not administered at a single testing session. Administering the test on different days would also be advantageous in reducing the effects of one testing on the other. These psychological considerations are judged to outweigh the administrative advantages which would arise if both tests in a pair were administered in a single session.

The decision to administer the two tests on different days affects the decision on whether the tests should be administered within a regular class session or administered to a student group assembled for the purpose. All tests involved in the study could be administered within a single class session of a course which meets for 40 minutes except for the TALENT Mathematics test. Administration as part of a regular classroom activity by the regular teacher would presumably be advantageous from the viewpoint of administrative convenience. Moreover, it would probably be relatively difficult to assemble a group drawn from a number of classes on two separate occasions, assuming that the testing is done during regular school hours. Although a probability sample

of students within school was tested in the NLS survey without undue difficulty, it does not seem essential for the equating study to avoid cluster sampling within each school. Provision should be made, however, for random selection of the class (or classes) to be tested.

The foregoing discussion implies that the ultimate sampling unit is a class of 12th grade students. More than one class could be tested within the same school. If this plan is followed, the following steps would be involved in the test administration:

(1) Ask the Chief State School Officer in each state represented in the sample of public schools to authorize our inviting the schools in that state to participate.

(2) Invite the Principal of each school represented in the sample to participate and to designate a Coordinator for the school.

(3) Ask the Coordinator to supply a list of all courses in which enrollment was composed predominantly of twelfth grade students, stating the times when each class meets and the name of the instructor. Only courses in which written tests could be given would be listed.

(4) Select at random one or more classes within each school.

(5) Supply instructions and test materials to the Coordinator for each class to be tested. Instructions would specify procedures for insuring that first and second testings for each student can be matched rigorously and efficiently. It will probably be necessary to use a system of numbered envelopes for each student. (The student would grid his school code and identifying number on each answer sheet to facilitate collating.)

(6) Tests would be administered by the classroom teacher in accordance with instructions provided.

Sample Design

Certain decisions about the plan for test administration clearly narrow the options available for efficient sample design. First, the decision to test in the classroom situation is likely to accentuate the effect of cluster sampling, especially if classes are differentiated for students in different programs and if tracking is used by a school. Second, the decision to administer only one sequence of a particular pair of tests in any one school

prevents matching of samples on the basis of class or school, which would be desirable from a sampling viewpoint.

Because this study is concerned with equating rather than with norming and especially because the equating method to be used takes advantage of the presumably high correlation between the tests to be equated, the foregoing difficulties should not prevent the development of an adequate sample design. Moreover, to the extent that the tests to be equated approximate strict parallelism, the characteristics of the sample should have no systematic effect on the line of relation between scores. The precision of equating at different score levels would be affected, however.

For the present study, stratification on the basis of school characteristics related to test performance is important for the following reasons: (1) to increase the efficiency of the sample, (2) to permit the substitution of a similar school for a school which declines to participate, and (3) to make the schools administering a given pair of tests in one order as similar as possible to the schools administering that pair of tests in the reverse order. Stratification may also be used to insure diversity in the sample even if the characteristic is not known to have a high correlation with test performance. A further consideration is that the analysis could be adapted to study the standard error of the samples considered as norms samples and thus provide concrete evidence on the extent to which the method is less efficient than simple random sampling of the same number of students. Although this analysis would be incidental to the main study, it would make a useful contribution to empirical evidence on sampling design.

For purposes of this study, the list of public and private secondary schools maintained by the United States Office of Education may be considered to provide an adequate basic sample source. Although the list is not a comprehensive list of all secondary schools, it does not seem necessary to augment it for this equating study. The tape record includes four items which are judged to be useful, directly or indirectly for stratification, as follows: (a) 12th grade enrollment, or 11th grade in schools for which 11 is the highest grade; (b) community in which school is located; (c) county in which school is located; and (d) state in which school is located.

The extensive analyses in Coleman et al. (1966) and in Mayeske et al. (1973) indicate that the geographical division in which the school is located and whether the school is located in a metropolitan or a nonmetropolitan region constitute useful bases for stratification. These two stratification bases have the advantage of being relatively meaningful. Moreover, the classification of schools with respect to these variables is unambiguous. It seems reasonable therefore to make an initial stratification on them. It is possible to use Statistics of Public Elementary and Secondary Day Schools, Fall 1972 (Foster, 1973) to obtain recent twelfth-grade enrollment figures for public schools by geographical region. Data for Montana and New Jersey were for earlier years. Data from the 1970 census provide information on public high school students (grades 9-12) which can be used to prorate enrollments into metropolitan and nonmetropolitan regions for each geographical division and also provide information on high school enrollments in private schools (U. S. Bureau of the Census, 1972 b). The following table shows estimated enrollments when public school students are classified on these two characteristics:

Geographical Division	Estimated Number of Students: Metropolitan	Nonmetropolitan
North Atlantic	548,602	153,723
Great Lakes and Plains	538,349	305,573
Southeast	279,972	311,966
West and Southwest	544,379	187,319

The total enrollment in public schools was 2,869,883.

Census data indicate that high school enrollment in private schools was about 10.8% of the public school enrollment (U. S. Bureau of the Census, 1972 a). For planning purposes, the private school twelfth grade enrollment may be taken as 309,947, and total public and private enrollment may be taken as 3,180,000.

If we combine data for nonmetropolitan regions outside the southeast, and divide the estimated enrollments by a suitable integer, we obtain the following table:

Public Schools	Estimated Enrollment	Number of Strata	Students per Stratum
Metropolitan			
North Atlantic	548,602	2	274,301
Great Lakes and Plains	538,349	2	219,174
Southeast	279,972	1	279,972
West and Southwest	544,379	2	272,190
Nonmetropolitan			
Southeast	311,966	1	311,966
All other regions	646,613	2	323,306
Private Schools			
All schools	309,947	1	309,947

This table indicates that the foregoing stratification would yield 11 principal strata which are similar in number of students. The difference in size, however, is sufficiently great to warrant differential weighting of the strata. In the study, the weights would be based on actual 12th grade enrollments obtained from the ELSEGIS file rather than on estimated enrollments.

The complete definition of the principal strata requries specification of ways of dichotomizing the four groups containing two strata. It is proposed that for the nonmetropolitan schools in the three regions combined, the stratification be based directly on 12th grade enrollment, divided so that the strata will be approximately equal. For metropolitan schools in each of the three regions having two strata, stratification would be based on the median years of education of persons 25 years or older and on the income level of families for each SMSA, as reported in the County and City Data Book, 1972.

In drawing the sample within each principal stratum, schools will be selected at random with a probability proportional to size, as determined by twelfth-grade enrollment obtained from the ELSEGIS Public and Non-Public school files.

The most difficult practical problems are likely to arise at the last stage, which calls for selecting one class, at random, within each school. Each school which agrees to participate in the study will be asked to supply a list of all classes composed mainly of 12th grade students which meet during the first full class period of a designated day of the school week and in which a

written test could be administered. The days of the week would be assigned at random in the invitations to participate. Classes would then be selected at random within school. Because the number of students actually tested would vary across schools, a weight would be assigned to each class to provide a uniform weighted sample size for all schools.

The determination of the appropriate sample size for this study is difficult because the sampling theory for equipercentile equating has not been developed. Fortunately, certain empirical results of the Anchor Test Study (Bianchini and Loret, 1972) are judged to be relevant to this question. The sampling design of the Anchor Test Study differs from the present study in that a more complex stratification plan was used, with oversampling of high and low socioeconomic status schools, with approximately half of the classes in each school assigned to each sequence of testing, and with pupil-weighting on the basis of sex, race, and IQ to adjust for absenteeism and unusable test data. In addition, the equipercentile equating method used in the Anchor Test study took advantage of a symmetrical equating design for seven tests to provide a broader base for equating each pair of tests. On the whole, the differences would be expected to result in greater precision for a given sample size in the Anchor Test study. Nevertheless, the Anchor Test Study results provide evidence on the order of magnitude of equating error arising when equipercentile equating is used in the counter-balanced design, with all members of the equating sample taking both tests to be equated.

Because the primary concern in this equating study is with sampling error in the vicinity of the mean, it is advantageous to bring together information from several parts of the Anchor Test Study. For this purpose, the following steps were taken:

(1) From the appropriate norms tables, the median of each type of reading score (Vocabulary, Comprehension, and Total) on the Metropolitan Achievement Tests (MAT) was obtained for grades 4, 5, and 6 separately.

(2) Using tables of equivalent scores, the score on each of the other six tests corresponding to each median MAT score was obtained.

(3) Tables 5-15, 5-16, and 5-17 of the Anchor Test Study Final Report present smoothed values of equating error for each of the six tests

equated to MAT. Using these tables, the average equating error, expressed in MAT raw score units, was obtained for the score on each test corresponding to the MAT median.

(4) The mean value of the six standard errors for each type of score was calculated.

(5) Using the norms tables for MAT, the standard deviation for each type of score at each grade level was estimated by taking one-half the difference between the 84th percentile and the 16th percentile.

(6) The mean standard error obtained in Step 4 was divided by the approximate standard deviation in Step 5 in order to express the standard errors in standard deviation units.

(7) Finally, the correlation between MAT and each of the tests to which it was equated was obtained, using the average coefficient for the two orders of administration. The median value for the six coefficients for each type of score at each grade level was calculated.

The foregoing steps yielded the results shown in Table 24. Error of equating expressed in standard deviation units has a median of .020 for the nine separate estimates. The range of the nine values is from .014 to .029. These findings indicate a very high degree of precision for the Anchor Test Study equatings.

The number of students having usable data in the Anchor Test Study was 134,855, as shown in Table 5-10 of the Final Report. When this figure is divided by 56, the number of separate equating units, and by 3, the number of grades, the average number of students who took a particular pair of tests in a particular order is 803.

The availability of the Anchor Test Study findings does not, of course, eliminate the need for judgment in determining the sample size for the study being planned. The Anchor Test Study design, as noted above, is more efficient than the design for this study. In addition, consideration must be given to the possibility that the correlation between the 12th grade tests will be smaller than those found in the Anchor Test Study, a factor which would also be expected to increase the error of equating. On the other hand, because this study is designed primarily to detect substantial changes in performance level, a standard error of equating equal to four percent of

Table 24

Standard Error of Equating, Standard Deviation of Metropolitan Achievement Test (MAT) Scores, and Median Correlation with MAT for Each Type of Reading Test Score in Grades 4, 5, and 6 (Source: Anchor Test Study)

Type of Reading Test Score	Grade	Mean of Standard Errors of Equating for 6 Tests ($\overline{SE}$)	Approximate Standard Deviation of MAT Scores* (SD_{MAT})	$\overline{SE}/SD_{MAT}$	Median Correlation with MAT
Vocabulary	4	.28	12.2	.023	.82
	5	.27	11.4	.024	.88
	6	.23	11.2	.021	.87
Comprehension	4	.32	11.1	.029	.82
	5	.20	10.2	.020	.83
	6	.20	10.4	.020	.85
Total	4	.43	22.4	.019	.88
	5	.35	21.2	.017	.91
	6	.30	20.8	.014	.91

*Estimated as one-half the difference between the 84th and 16th percentiles.

the standard deviation should be adequate. The error would, of course, be larger for the subgroup comparisons involved in evaluating test equivalence.

On the whole, the foregoing considerations indicate that the sample for each of the eight groups required to provide two orders of administration of each of the four test pairs should include two schools (and, therefore, two classes) from each of the eleven principal strata. Assuming 25 students per class, each of the eight samples would include 550 students. Classes within stratum would be assigned at random to each half-sample for calculating errors of equating and of norms. In all, usable data for 176 schools will be needed. In drawing the sample of schools, eight schools will be selected from each principal stratum as possible replacements for schools in the initial sample which do not agree to participate.

Data Analysis

Because the study design and data analysis are so closely connected in a study of this kind, the following statement of analysis steps is essentially a recapitulation.

Editing. Only students who took both tests under defined conditions will be included in the analysis sample. Data will also be edited on the basis of teacher's irregularity reports, and answer sheets which have two or fewer items answered will be excluded.

Weighting. A weight for each student will be determined based on the following two factors: (1) one-eleventh divided by the proportion of all 12th grade students who belong to his or her principal stratum; and (2) 25 divided by the number of usable cases in his or her class. The same weight will be used in all phases of the analysis for a given student. Under this weighting plan, all members of a particular class will have the same weight.

Equating. Equating results for the four pairs of tests will be obtained by averaging the equipercentile equating results for the two orders of administration for the following groups of students for each test pair:

(a) All students;

(b) Boys and girls separately;

(c) Students dichotomized on the basis of parents' education;

(d) Students dichotomized on the basis of type of school program (Academic vs. all others);

(e) Students dichotomized on the basis of self-reported high school grades;

(f) For mathematics tests only, students dichotomized on the basis of amount of high school study of mathematics; and

(g) Each of the eleven half-samples of the group taking each test pair, for use in determining equating error.

In the event that results for one or more of the subgroup equatings done to aid in evaluating equivalence yield ambiguous results, equating errors for each of the two subgroups will be developed using balanced half-sample replication.

Error of Norms. Although this study is not a norms study, the balanced half-sample replication design will permit the evaluation of the sampling error of the mean score on the NLS Reading test and the NLS Mathematics test, when taken as the first test in a pair. This sampling error will be useful in evaluating the results obtained when the corresponding EEOS or TALENT test scores obtained in the samples which took them as the first test of the pair are expressed in NLS units. The line of relation obtained in the study would be used for this conversion. In addition, the results will provide useful information on the size of the error arising when classes rather than individual students or schools are sampled at the high school level.

Practice Effect. By converting scores for tests other than NLS to the appropriate NLS scale, it will be possible to obtain means and standard deviations for all initial and final tests on the NLS scale. This analysis should throw some light on whether or not practice effects operate symmetrically for the two members of a test pair.

Parallelism. The correlation between the two tests in each pair can be evaluated in relation to the reliability of scores for each test in the pair. For this analysis, Kuder-Richardson Formula No. 20 reliabilities will be determined for each test, using the sample which took that test first.

Application of Conversions to Survey Data. Assuming that the outcome of the equivalency comparisons warrants the conversion of scores of tests in one survey to those in another, NLS data files will be utilized to convert scores

for NLS tests to the appropriate scores in each of the other two surveys and to provide the same statistics for corresponding groups as have been provided in the earlier surveys.

For maximum precision of the comparisons, certain additional tabulations of TALENT, EEOS, and NLS data should be performed. Obtaining NLS distributions, means, and standard deviations based on weighted N's, using the weights devised for the test participants rather than those for all participants would be useful. These statistics should be obtained for the total group having test scores, for male and female students separately, and for black and white students separately. Moreover, to permit comparison with EEOS statistics, tabulations based on public school students only should be obtained. For EEOS, separate tabulations for male and female students should be obtained, provided that this can be done without excessive cost. Finally, the Project TALENT Data Bank should be utilized to obtain distributions, means, and standard deviations for public school students only and for the total group. For both of these groups, data should be analyzed separately by sex, and for the two sexes combined. The TALENT results for reading comprehension should be compared with the results of a 1960 to 1970 comparison based on administration of the TALENT test in 1970 (Flanagan and Jung, 1971). The stratifications by sex and race do not, of course, exhaust the possible subgroups which might be examined to provide trend data for significant student groups.

REFERENCES

Angoff, W. H. Can useful general-purpose equivalency tables be prepared for different college admissions tests? In A. Anastasi (Ed.) Testing problems in perspective. Washington: American Council on Education, 1966. Pp. 251-264.

Angoff, W. H. Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.) Educational Measurement, Second Edition. Washington: American Council on Education, 1971. Pp. 508-600.

Bianchini, J. C. & Loret, P. G. Anchor Test Study: Final report. Princeton: Educational Testing Service, 1972.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. Equality of Educational Opportunity. Washington: Government Printing Office, 1966.

Cooperative Test Division, Educational Testing Service. Cooperative Sequential Tests of Educational Progress: Technical Report. Princeton: Educational Testing Service, 1957.

Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorham, W. A., Orr, D. B., and Goldberg, I. Designing the Study. (Technical Report to the U. S. Office of Education, Cooperative Research Project No. 566.) Pittsburgh: Project TALENT Office, University of Pittsburgh, 1960. (ERIC No. 002198)

Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorham, W. A., Orr, D. B., & Goldberg, I. Design for a Study of American Youth. Boston: Houghton Mifflin, 1962.

Flanagan, J. C., Davis, F. B., Dailey, J. T., Shaycoft, M. F., Orr, D. B., Goldberg, I., & Neyman, F. A., Jr. The American High School Student. (Final report to the U. S. Office of Education Cooperative Research Project No. 635.) Pittsburgh: Project TALENT Office, University of Pittsburgh, 1964.

Flanagan, J. C., and Jung, S. M., Progress in Education: A Sample Survey (1960-1970). Palo Alto, California. American Institutes for Research, 1971.

Foster, B. J. Statistics of Public Elementary and Secondary Day Schools, Fall 1972. DHEW Publication No. (OE) 73-11402. Washington: Government Printing Office, 1973.

Hilton, T. L., Rhett, H., Broudy, I. L., Bower, C., Carter, M. M., Creech, F. R., & Echternacht, G. J. The Base-Year Survey of the National Longitudinal Study of the High School Class of 1972. Final Report. Princeton: Educational Testing Service, 1973.

Horvitz, D. G., Mason, K., Bayless, D., Jackson, D., Hunt, N., and Koch, D. Design of a Longitudinal Survey of Secondary School Seniors. (2 Volumes) Research Triangle Park, North Carolina: Research Triangle Institute, 1972.

Kish, L. & Frankel, M. R. Balanced repeated replication for standard errors. Journal of the American Statistical Association. 1970, 65, 1071-1094.

Kish, L. & Frankel, M. R. Inference from complex samples. Journal of the Royal Statistical Society, Series B. 1974, 36, 1-22.

Lord, F. M. Notes on comparable scales for test scores. Research Bulletin 50-48. Princeton: Educational Testing Service, 1950.

Mayeske, G. W., Okada, T., Beaton, A. E., Jr., Cohen, W. M., & Wisler, C. E. A study of the achievement of our nation's schools. DHEW Publication No. (OE) 72-131. Washington: Government Printing Office, 1973.

McCarthy, P. J. Pseudo-replication: Half-Samples. Review of the International Statistical Institute. 1969, 37, 239-264.

Plackett, R. L. & Burman, J. P. The design of optimum multifactorial experiments. Biometrika, 1943-1945, 33, 305-325.

Project TALENT, The Project TALENT Data Bank: A Handbook. Palo Alto, California: American Institutes for Research, 1972.

Shaycoft, M. F.; Dailey, J. T.; Orr, D. B.; Nayman, C. A., Jr.; and Sherman, S. E. Studies of a Complete Age Group-Age 15. (Technical Report to the U. S. Office of Education, Cooperative Research Project No. 226.) Pittsburgh: Project TALENT Office, University of Pittsburgh, 1963.

U. S. Bureau of the Census, Census of Population: 1970. General Social and Economic Characteristics. Final Report PC (1) C1 United States Summary. Washington: Government Printing Office, 1972a.

U. S. Bureau of the Census, Census of Population: 1970. General Social and Economic Characteristics. Final Report PC (1) C2 (51 volumes). Washington: Government Printing Office, 1972b.

U. S. Bureau of the Census, Public Use Samples of Basic Records from the 1970 Census: Description and Technical Documentation. Washington: 1972c.

U. S. Bureau of the Census, Statistical Abstract of the United States, 1973. (94th edition) Washington: Government Printing Office, 1973.

WESTAT, INC., Sample Design for the Selection of a Sample of Schools with Twelfth-Graders for a Longitudinal Study. Rockville, Maryland: WESTAT, INC., 1972